# MECHANICAL COMPONENT FAILURE PROGNOSIS STUDY

Hans K. Ziebarth, et al

Garrett Corporation

Prepared for:

Army Air Mobility Research and Development Laboratory

June 1973

Unclassified
Security Classification

AD 771033

## DOCUMENT CONTROL DATA - R & D

*(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)*

1. ORIGINATING ACTIVITY *(Corporate author)*
AiResearch Manufacturing Company, a division of The Garrett Corporation
Torrance, California

2a. REPORT SECURITY CLASSIFICATION
Unclassified

2b. GROUP

3. REPORT TITLE
MECHANICAL COMPONENT FAILURE PROGNOSIS STUDY

4. DESCRIPTIVE NOTES *(Type of report and inclusive dates)*
Final Report

5. AUTHOR(S) *(First name, middle initial, last name)*
Hans K. Ziebarth
Jee-Da Chang
Joseph Kukel

| 6. REPORT DATE | 7a. TOTAL NO. OF PAGES | 7b. NO. OF REFS |
|---|---|---|
| June 1973 | ~~99~~ 102 | 33 |

| 8a. CONTRACT OR GRANT NO. | 9a. ORIGINATOR'S REPORT NUMBER(S) |
|---|---|
| DAAJ02-71-C-0049 | |
| b. PROJECT NO. Task 1F162203A43405 | USAAMRDL Technical Report 73-26 |
| c. | 9b. OTHER REPORT NO(S) *(Any other numbers that may be assigned this report)* AiResearch Report No. 72-8745 |
| d. | |

10. DISTRIBUTION STATEMENT
Approved for public release; distribution unlimited.

| 11. SUPPLEMENTARY NOTES | 12. SPONSORING MILITARY ACTIVITY |
|---|---|
| | Eustis Directorate, U. S. Army Air Mobility Research and Development Laboratory, Fort Eustis, Virginia |

13. ABSTRACT

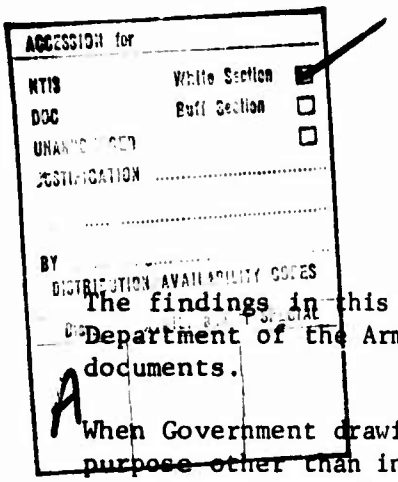
This report presents the results of a study of failure prognosis of mechanical components of propulsive transmissions and gearboxes of Army aircraft. The step from on-line determination of component current mechanical condition by inferential, or diagnostic, methods to condition prognosis appears feasible if diagnostic techniques of the signature-composite interpretive type are used and diagnostic information is blended with inputs from the areas of component design theory, tribology, system abnormal physiology, and life predictive theory. For current component status determination, use of dynamic and lubricant-carried particle content signatures was found to be of greatest value. Use of the composite exceedance method in the dynamic data interpretation area and use of statistical methods recently developed in fluid-borne particle metrology in the lubricant particle content area would be indicated for attaining prognostic objectives.



DD FORM 1 NOV 68 1473 REPLACES DD FORM 1473, 1 JAN 64, WHICH IS OBSOLETE FOR ARMY USE.

Unclassified
Security Classification

| 14. KEY WORDS | LINK A | | LINK B | | LINK C | |
|---|---|---|---|---|---|---|
| | ROLE | WT | ROLE | WT | ROLE | WT |
| Mechanical Component Condition Prognosis | | | | | | |
| Diagnostic Prognostic Systems | | | | | | |
| Helicopter Transmissions and Gearboxes | | | | | | |
| Life Prediction | | | | | | |
| Failure Prediction | | | | | | |
| Reliability | | | | | | |
| Mechanical Component Failure Modes | | | | | | |




## DISCLAIMERS

The findings in this report are not to be construed as an official Department of the Army position unless so designated by other authorized documents.

When Government drawings, specifications, or other data are used for any purpose other than in connection with a definitely related Government procurement operation, the United States Government thereby incurs no responsibility nor any obligation whatsoever; and the fact that the Government may have formulated, furnished, or in any way supplied the said drawings, specifications, or other data is not to be regarded by implication or otherwise as in any manner licensing the holder or any other person or corporation, or conveying any rights or permission, to manufacture, use, or sell any patented invention that may in any way be related thereto.

Trade names cited in this report do not constitute an official endorsement or approval of the use of such commercial hardware or software.

## DISPOSITION INSTRUCTIONS

Destroy this report when no longer needed. Do not return it to the originator.

DEPARTMENT OF THE ARMY
U. S. ARMY AIR MOBILITY RESEARCH & DEVELOPMENT LABORATORY
EUSTIS DIRECTORATE
FORT EUSTIS, VIRGINIA 23604

This report was prepared by AiResearch Manufacturing Company, a division of the Garrett Corporation, under the terms of Contract DAAJ02-71-C-0049. It presents the findings of a 12-month study of the definition of a diagnostic/prognostic methodology for mechanical components of Army aircraft propulsive transmissions and gearboxes.

The objective of this contractual effort was to analyze the feasibility of predicting the useful life remaining, or rate of failure progression, for gears and bearings that have experienced an incipient failure and to recommend future research efforts needed to realize that prediction capability. Prognostic principles being used in other technology areas were to be examined to determine their possible use for the subject purpose. A comprehensive evaluation of the following monitor techniques was performed:

a. Vibration monitoring

b. Composite exceedance monitoring

c. Acoustic emission monitoring

d. Nondestructive testing

e. Lubricant particle content analysis

The conclusions and recommendations contained herein are generally concurred in by this Directorate. The study appears to have adequately fulfilled the major objectives.

Technical direction for this contractual effort was provided by Mr. Roger J. Hunthausen, Military Operations Technology Division.

TASK 1F162203A43405
Contract DAAJ02-71-C-0049
USAAMRDL TECHNICAL REPORT 73-26
June 1973

# MECHANICAL COMPONENT FAILURE PROGNOSIS STUDY

Final Report

AiResearch Manufacturing Company Report 72-8745

By

H. K. Ziebarth
Dr. J. D. Chang
Dr. J. Kukel

Prepared by

AiResearch Manufacturing Company
a Division of
The Garrett Corporation
Los Angeles, California

for

EUSTIS DIRECTORATE
U.S. ARMY AIR MOBILITY RESEARCH AND DEVELOPMENT LABORATORY
FORT EUSTIS, VIRGINIA

## SUMMARY

This report presents the findings of a 12-month study of a condition-prognostic methodology for mechanical components of Army aircraft propulsive transmissions and gearboxes.

The following goals were achieved during the course of the study:

1. Prognostically relevant inputs from areas other than diagnostics were identified.

2. Structure-borne dynamic and lubricant particle content signatures were determined to represent the most fruitful prognostic parameters. Advanced interpretation methods for these signatures were defined.

3. From a prognostic viewpoint, a comprehensive evaluation of monitoring techniques was performed. These included:

   a. Vibration monitoring

   b. Composite exceedance analysis

   c. Acoustic emission monitoring

   d. Nondestructive testing (NDT)

   e. Lubricant particle content analysis

4. Prognostic prediction methodologies, such as those based on single-point, multipoint, and nonquantifiable time history data, were evaluated.

5. Knowledge of concrete life-prediction mechanisms was found to be lacking. The current electronic technology level, however, appears adequate to implement data processing and computational requirements of envisaged approaches in an operational, on-line manner.

Although substantial experimental and analytical effort along the guidelines laid down by this study must still be expended to provide a basis for a mechanical component condition forecasting epistemology, the study established confidence that a predictive model can be achieved.

## FOREWORD

This report presents the results of a study conducted by AiResearch Manufacturing Company, a division of The Garrett Corporation, for the Eustis Directorate, U. S. Army Air Mobility Research and Development Laboratory, under Contract DAAJ02-71-C-0049 (Task 1F162203A43405). The USAAMRDL technical monitor was Mr. Meyer B. Salomonsky.

Principal investigators for the applied systems research group at AiResearch were H. K. Ziebarth and Dr. J. D. Chang. Dr. J. Kukel was project manager for the study.

Contributors included Dr. H. Friedericy, R. P. Chen, and R. Bhika of the stress and bearing groups of the preliminary design department, and W. J. Harris of the applied systems research group.

## TABLE OF CONTENTS


| | Page |
|---|---|
| SUMMARY | iii |
| FOREWORD | v |
| LIST OF ILLUSTRATIONS | ix |
| LIST OF TABLES | x |
| LIST OF SYMBOLS | xi |
| INTRODUCTION | 1 |
| METHODOLOGICAL CONSIDERATIONS | 2 |
| CURRENT STATUS VERIFICATION | 5 |
| Operational Instrumentation | 5 |
| Diagnostic Methods | 5 |
| Nondestructive Test Methods | 38 |
| PROGNOSTIC IMPLICATIONS OF TRANSMISSION MECHANICAL DESIGN CHARACTERISTICS | 57 |
| Modular Transmission Design | 57 |
| Functional Redundancy of Components of Life-Critical Subassemblies | 57 |
| Basic Load Concentration | 57 |
| Specific Load Concentration | 58 |
| Lubrication Mode | 60 |
| Tribological Conditions | 60 |
| FAILURE MODES AND PATHS | 62 |
| RELIABILITY STATISTICAL LIFE PREDICTION | 65 |
| Bearing Fatigue Life Modeling | 65 |
| PROGNOSTIC METHODOLOGY | 74 |
| Prognostic Analysis of Single-Point Time History Data | 74 |
| Analysis of Multipoint Time History Data | 75 |
| Analysis of Nonquantifiable Time History Data | 76 |
| CONCLUSIONS | 77 |
| RECOMMENDATIONS | 78 |
| RECOMMENDED FOLLOW-ON PRIORITIES | 84 |

## CONTENTS (continued)


| | Page |
|---|---|
| LITERATURE CITED ..... | 85 |
| DISTRIBUTION ..... | 88 |

## LIST OF ILLUSTRATIONS


| Figure | | Page |
|---|---|---|
| 1 | Failure Prognosis: Matrix of Supporting Disciplines and Methods | 3 |
| 2 | A Typical Vibration Signal from a Gear Pump | 8 |
| 3 | Power Spectral Density Distribution of a Typical Narrow-Band Gaussian Signal | 9 |
| 4 | Characteristics of a Composite Exceedance Curve | 11 |
| 5 | Comparison of an Actual Exceedance Curve With a Reference Exceedance Band | 13 |
| 6 | Example of Peak-Counting Procedures | 16 |
| 7 | Typical Acoustic Emission Signals | 19 |
| 8 | Typical Acoustic Emission Analysis Setup | 21 |
| 9 | Mechanical and Acoustic Emission Fatigue Test Results | 23 |
| 10 | Typical Pressure Vessel Test Acoustic Emission Plot | 25 |
| 11 | Sample Result of Stress Corrosion Tests | 26 |
| 12 | Acoustic Emission Generated by Weld Faults | 27 |
| 13 | Detection of Incipient Bearing Failure by Acoustic Emission/Shock Pulse From a High-Frequency Accelerometer Output | 29 |
| 14 | Detection of Incipient Bearing Failure by Acoustic Emission/Shock Pulse From a Lightly Damped Sonic Accelerometer Output | 30 |
| 15 | Contaminant Concentrations for Gravimetric Tolerance Tests | 33 |
| 16 | Fluid Cleanliness Level Requirements | 33 |
| 17 | Change of Lubricant Light Transmittance as a Function of Flight Hours | 37 |
| 18 | Change of Lubricant Light Transmittance as a Result of Operation at Elevated Temperature | 37 |

## LIST OF TABLES


| Table | | Page |
|---|---|---|
| I | Peak Counting Results ..... | 16 |
| II | Wear Process Equation Exponent Values ..... | 31 |
| III | Lubricant-Contained Prognostic Signature Analysis Methods ..... | 34 |
| IV | Characteristics of the NDT Techniques ..... | 40 |
| V | Values for Constants m and n..... | 68 |

## LIST OF SYMBOLS

### Dynamic Signature Analysis

| | |
|---|---|
| $C$ | Count function |
| $f$ | Frequency |
| $m$ | Number of speed levels |
| $N$ | Number of peaks per second |
| $N_o$ | Constant related to power spectral density |
| $P$ | Utilization factor |
| $T$ | Duration of test run |
| $t$ | Time instant |
| $z$ | Signal |
| $\xi$ | Amplitude level |
| $\sigma$ | Total power content |
| $\Phi$ | Spectral density function $\Phi_{zz}(f)$: auto power spectral density function of signal z |
| $\Omega$ | Specific speed level |

### Acoustic Emission Signature Analysis

| | |
|---|---|
| $B$ | Proportionality constant |
| $N$ | Total emission count |
| $\epsilon$ | Strain level |
| $\epsilon_o$ | Threshold strain level |
| $\omega$ | Number of cracked fibers |

### Wear Particle Generation Monitoring

| | |
|---|---|
| $d$ | Distance |
| $m, n, p$ | Exponents signifying wear mode |

## LIST OF SYMBOLS (Continued)

| | |
|---|---|
| U | Velocity |
| V | Wear volume |
| W | Load |
| a,c,d,f | Subscripts signifying wear mode |

Lubricant Physical-Chemical Condition

| | |
|---|---|
| $\Omega$ | Resistivity |

Specific Load Concentration

| | |
|---|---|
| d | Operating pitch diameter of pinion |
| F | Net face width |
| $n_p$ | Pinion speed |
| $P_{ac}$ | Allowable power transmitted |

Bearing Fatigue Life Modeling

| | |
|---|---|
| A,B | Bearing curvature characteristics |
| a | Pressure ellipse major axis; derived quantity |
| b | Pressure ellipse minor axis; derived quantity |
| c,e,h | Exponents reflecting material properties |
| E | Young's modulus; expected damage |
| D | Total damage |
| k | Factor reflecting geometric characteristics of contact area |
| $\ell$ | Length of rolling path |
| $\ell_e$ | Roller length |
| m,n | Functions of geometric characteristics |
| N | Number of $10^6$ cycles |

## LIST OF SYMBOLS (Continued)

| | |
|---|---|
| $n$ | Number of stress peaks; number of applied stress cycles |
| $P$ | Total radial load; probability density |
| $R$ | Minimum radius of curvature |
| $R'$ | Maximum radius of curvature |
| $S$ | Survival probability; design requirements parameter; equivalent tensile stress |
| $s$ | Stress level |
| $x$ | Angle between planes of curvatures |
| $Z$ | Orthogonal shear stress depth |
| $\beta, \varkappa$ | Constants |
| $\mu$ | Sliding coefficients of friction |
| $\rho$ | Surface pressure |
| $\tau$ | Orthogonal shear stress |
| $\nu$ | Poisson's ratio |

# INTRODUCTION

The determination of the mechanical condition of rotary-wing aircraft propulsion systems by operational diagnostic methods has been the objective of a series of recent programs. These methods, used with the system in the installed, operative condition, determine the current condition of the propulsion system including transmissions and gearboxes.

Results of these programs, together with urgent user requests, raised the question of the feasibility of progressing at this time from diagnosis to prognosis, or from present mechanical status determination to prediction of a system's future state based on knowledge of its past and present history.

During the past five years, the art of condition diagnosis of mechanical systems has progressed as rapidly as the understanding of the mechanisms involved in the life expenditure process of the machine elements such as gears and bearings. In the same period, significant advances have been made in the application of probabilistic methods to the design life theory of mechanical systems, and in the refinement of methods of this type as part of reliability theory. The coincidental simultaneous pursuits in the fields of mechanical system diagnosis, tribology, and prediction theory have greatly enhanced the chances for success in mechanical component failure prognosis.

Compared with life prediction in the frame of reliability studies, near real-time prediction of the future state of mechanical systems is a widely unexplored area. It could not be expected, therefore, that the study program reported here could establish prognostic procedures ready for application, for example, to such complex and critical problems as helicopter transmission component failure prediction. The present study, however, appears to have fulfilled its major objective, which was to provide a basis for future specific efforts to establish such procedures.

## METHODOLOGICAL CONSIDERATIONS

Elementary prognosis consists of projecting a mechanical system's future condition by observing the rate of change of one or more condition-reflecting parameters toward an empirically preestablished limit that signals the end of useful system life. An example of this type of prognosis would be the determination of transmission time to failure by trending of its wear particle generation and overall vibration characteristics only. This relatively established type of failure prediction reduces the life expenditure of a multicomponent, complex mechanical system to a process reflected in two basic, measurable signatures. Obviously, if a severe failure condition could be initiated by fatigue of a single critical bearing with a short remaining life margin, the practical value of this prognostic approach would be limited.

Even if a larger number of status-reflecting signatures were used for diagnosis, mechanical condition monitoring represents only one of many available information areas for constructing a prognostic methodology. Figure 1 presents this situation schematically. Information from four areas other than current status verification can be used for failure prognosis. Those stemming from the areas of mechanical design theory and failure phenomenology (combined component degradation and failure modes) appear to be the most significant. The static and dynamic stress levels assumed for the selection of critical bearings are examples of the usefulness of inputs from the mechanical design theory area. The time span between lapse of life or fatigue and functional failure is primarily a function of this composite stress level. Neglect of this design factor, therefore, could lead to failure to realize the need for a prognostic approach with extremely high response capability to cover a transmission failure case initiated by components of small remaining life margin. For this type of rapid failure, exclusive use of current status information on vibration and particle generation signatures would eliminate prognostic possibilities, since the usefulness of these signatures depends on a relatively progressed change of the failing component's exterior conditions. Of course, current status information should not be limited to those two types of information, since the recent development of methods such as acoustic emission or active ultrasonic signature monitoring could allow future detection of an incipient component structural degradation significantly earlier than by the more established methods, depending on distinct component physical condition changes. Even in that case, however, use of inputs from the mechanical design theory area may be valuable to initiate or emphasize certain prognostic routines during a specific sector of component life.

The engineering disciplines categorized as tribomechanics and tribochemistry, concerned with the clarification of the physicochemical processes in lubricated rolling contact mechanisms, contain a great deal of data relevant to mechanical component failure prediction. Since tribological conditions under which bearing and gear components operate are determined to a large degree by the system's design characteristics, the tribological data and the design theory tributaries in the methodology structure shown in Figure 1 have been ranked equally.

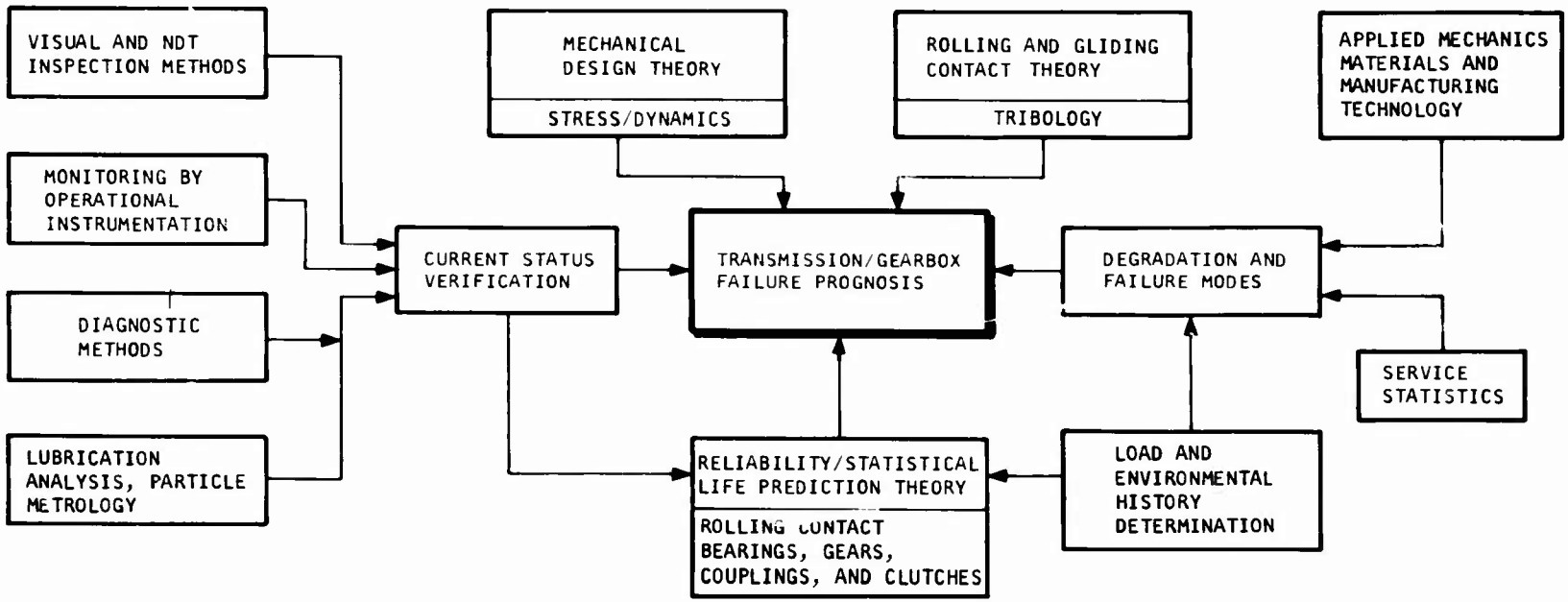



Figure 1. Failure Prognosis: Matrix of Supporting Disciplines and Methods.

Component degradation and failure mode data relate to failure prediction in two ways: (1) the character of a mechanical component degradation influences the selection of the type of status verification technique needed to detect and follow its progress; (2) knowledge of the paths along which failure conditions progress or proliferate can define optimum points for the interception of a failure process in view of signature character, signal strength, and instrumental convenience.

Probabilistic life prediction, the domain of reliability theories, can provide guidelines for prognostic concentration on mechanical system components that rank high both in failure frequency criticality and proliferation potential.

As shown in Figure 1, by using relevant state-of-the-art prognostic information, study activities were planned to progress from a systematic evaluation of the major source areas to a preliminary definition of how this information concept could be used to define prognostic methods.

## CURRENT STATUS VERIFICATION

### OPERATIONAL INSTRUMENTATION

Current system status can be verified by (1) data from operational instrumentation (such as lubricant bulk temperature in a typical helicopter transmission), (2) visual inspection or nondestructive test techniques, and (3) application of diagnostic methods that lead to inferential statements on component mechanical condition based on a pattern of individually observed signatures. The objective of this study was to determine the feasibility of failure prognosis based on data obtained from an operational, installed system; therefore, only current status verification methods of the first and third types are of primary interest, with inspection and nondestructive test techniques furnishing only complementary data. For these reasons, the state of the art of nondestructive test techniques was reexamined in this study to determine if certain of these techniques, with further development, might have potential usefulness for flight-line and possibly in-flight use; this question was later answered affirmatively. Information obtainable from operational instrumentations presently provided (for instance, on lubricant bulk temperature) has only limited prognostically useful content. If transmission heat rejection to lubricant were measurable by instrumenting both for inlet and outlet temperature, this information content would be increased, but not to a level warranting the additional cost involved. This is primarily because the majority of transmission component failure modes would result in significant changes of transmission heat rejection only if relatively far progressed.

The extension of operational instrumentations to include vibration sensors makes this a diagnostic method. It is discussed first because of its primary prognostic significance.

### DIAGNOSTIC METHODS

For a mechanical system such as a transmission, parametrical/functional diagnosis (i.e., condition determination on the basis of performance or functional parameters) offers only limited possibilities, since speed or torque relationships would be affected only by mechanical condition changes of an already progressed type. Additionally, measurement of input and output torque, for instance, in an effort to detect transmission efficiency reductions as a result of mechanical component condition degradations, not only would be difficult to implement but, since it would provide only gross information on overall mechanical condition changes, would be of limited prognostic value in the sense of not allowing identification of failing components.

Vibration and wear particle generation monitoring traditionally have ranked high among diagnostic methods used for the current status verification of transmissions and gearboxes. Comparing these two methods, vibration monitoring in its widest sense (i.e., including the various more recently developed techniques addressing signatures reflecting the mechanical condition

of a component in its dynamic state) has the higher resolution capability in distinguishing between the dynamic emanations from individual components or component elements inside a mechanical system. Wear particle generation monitoring methods lack this capability unless the system design is of the modular type or trace-metal impregnation of components is used. If modular design, or subassembly compartmentalization, is carried to the point of providing individual lubricant run-offs from critical bearings in which particle sensors can be installed, wear particle generation monitoring, because of its real-time information output, in view of prognostic objectives could prove to be superior to dynamic monitoring techniques.

## Dynamic Signature Analysis

Of major importance among techniques indirectly providing information on system component mechanical condition ( i.e., by other than visual or NDT type inspection methods) are those based on analysis of dynamic signatures such as bearing and gear vibrations or shock resonance pulses, reflecting changes of the physical condition of components as a result of normal life expenditure, accelerated wear, or incipient failure.

Widely explored and essentially established techniques for the status determination of mechanical systems include broadband amplitude analysis, deterministic spectrum correlation, statistical learning, and pattern recognition, extending over the sonic as well as the ultrasonic vibration signature regime. Commercially available hardware (SKF's Shock Pulse Measuring System MEPA 10A, Reference 1) is the result of the development of a technique using dynamic signature based on shock pulse phenomena. Applications of this technique to status determination of aircraft propulsion or propulsive transmission systems presently appear to be only exploratory in nature (see Figures 9 and 10 taken from Reference 2). Early presentations of the shock pulse method (Reference 3) emphasized its capability to study the progress of bearing raceway or rolling element flaking from first occurrence to lapse of additional life. Availability of the shock pulse method also may help in more clearly defining minimum and additional life, thus overcoming the present conceptual impasse that classically defines bearing life as extending from zero to life.

Of particular interest among dynamic methods is the composite exceedance method (recently developed by AiResearch for high-confidence production checkout of the drive subsystem of the Spartan missile secondary power generation system). The method furnishes a composite dynamic characteristic of the subsystem modified by mechanical condition changes. The manner in which this characteristic is changed from its normal shape can provide clues to the type of mechanical condition change that has occurred, thus helping to determine if it is of a life-critical type. Compared to vibration monitoring techniques of types other than broadband, particularly the deterministic spectrum correlation techniques, composite exceedance also points out the existence of mechanical irregularities of as-yet unidentified types. The technique, in generating time-sequential system total dynamic characteristics, permits monitoring of the progress or proliferation of an

originally localized mechanical abnormality. From the standpoint of implementation, the computational effort needed for the composite exceedance method considerably exceeds that for other dynamic analysis methods, with the deterministic spectrum correlation method approximately in the middle.

Following is a discussion of the principle of operation, special application considerations, methods of establishing reference composite exceedance curves, and generation of composite exceedance curves during operation.

## Principle of Operation

This method also uses one of the side effects, i.e., vibration signature, of a mechanical component to monitor the health status of that component. It differs, however, from the conventional broadband vibration monitoring and the spectrum analysis of vibration signals in that it uses an indirect property extracted from the vibration signal instead of the direct properties such as amplitude and frequency of the vibration signal.

The method makes use of the well-known fact that a mechanical component or a simple mechanical assembly or system generates a semiperiodic vibration signal (Figure 2) when operated under a given load and speed condition. The signal generated is semiperiodic because it deviates from a simple harmonic vibration in three respects. First, the amplitude varies from time to time following a slow varying envelope (dotted lines in Figure 2). Second, the periodicity or the frequency of the vibration signal changes over a range of values during normal operation under a constant condition (as evidenced by the different period $T_1$ and $T_2$ of the same signal at two time instants). Third, phase abruptions also occur frequently during a normal constant condition run (such as circle "A" in Figure 2). Signals of this type are called narrow-band random signals by electronic communications engineers and can be adequately described by a known statistical model called narrow-band Gaussian random process. A narrow-band Gaussian random process also can be characterized in the frequency domain in terms of frequency spectrum plots or power spectral density distributions as a closely clustered bunch of frequencies around a larger centroid frequency $f_o$, as shown in Figure 3. This random process is such a fully explored area that mathematical expressions of the process are available, and pertinent properties of this type of random signal can be derived from these mathematical equations and expressed in analytical forms.

One of the pertinent properties that finds wide application to statistical structural dynamics is the composite exceedance curve, i.e., the number of cyclic peaks above or equal to a given amplitude level within a given time. This quantity has been theoretically derived for a stationary narrow-band Gaussian random vibration of zero mean amplitude to be the following:

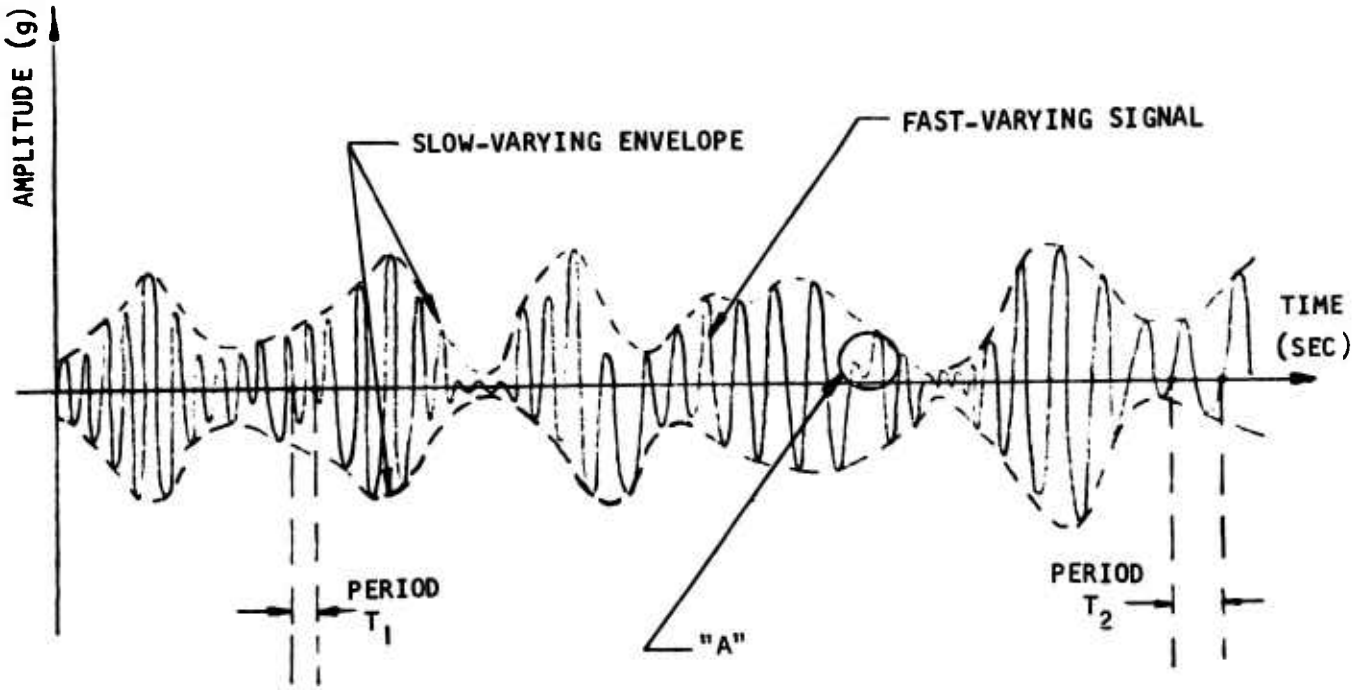



Figure 2. A Typical Vibration Signal From a Gear Pump.

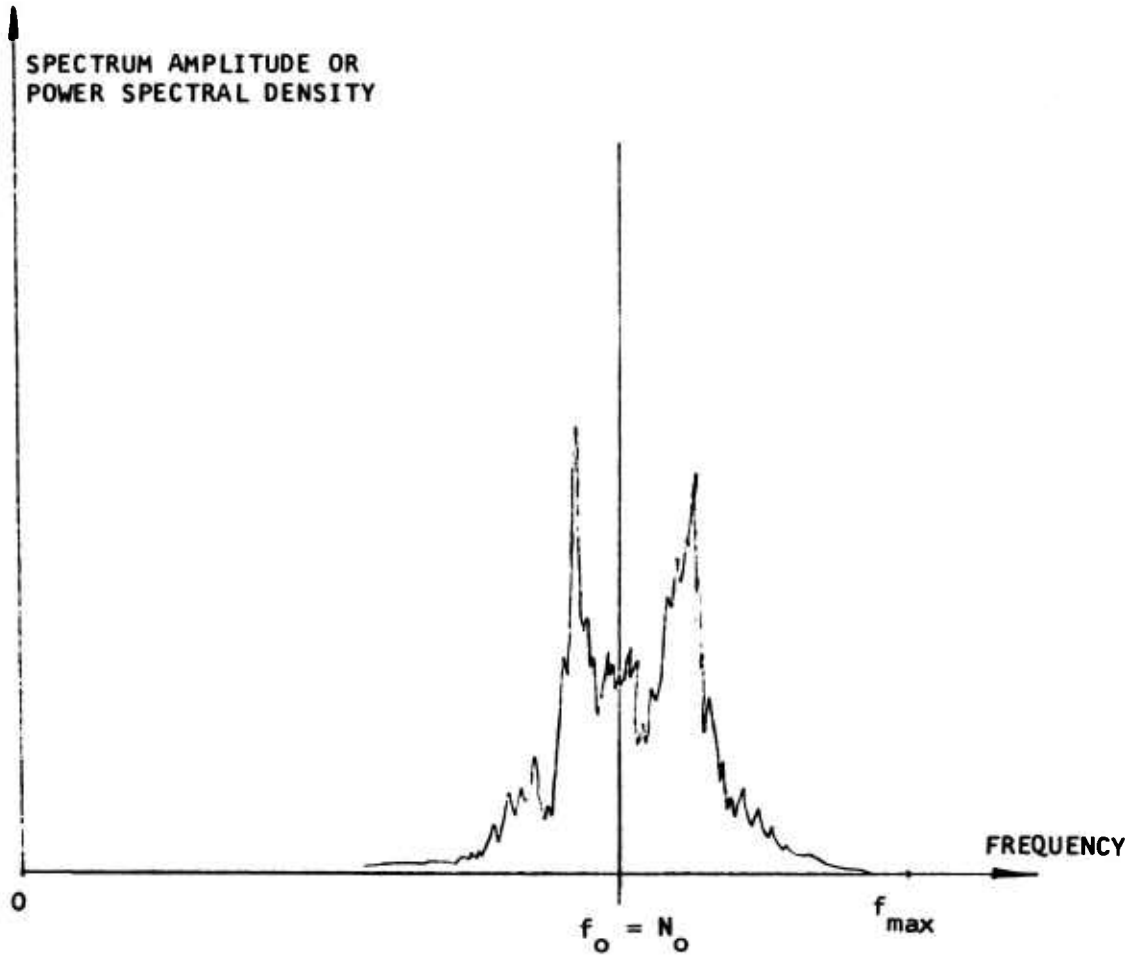



Figure 3. Power Spectral Density Distribution of a Typical Narrow-Band Gaussian Signal.

$$N(\xi) = \frac{N_o}{2\pi} e^{-\frac{\xi^2}{2\sigma^2}} \qquad (1)$$

where $N$ = number of peaks per second that exceed or equal the amplitude level $\xi$ and is a function of $\xi$

$\xi$ = amplitude level of the vibration signal and is the independent variable in Equation (1)

$\sigma$ = constant that is determined from the power spectral density of the vibration signal

$N_o$ = another constant that is determined also from the power spectral density of the vibration signal

For a given power spectral density distribution $\Phi_{\dot{Z}\dot{Z}}(f)$, the constant $\sigma$ is the total power content of the vibration signal and can be computed from Equation (2).

$$\sigma^2 = \int_o^{f_{max}} \Phi_{\dot{Z}\dot{Z}}(f)\,df \qquad (2)$$

Similarly, the constant $N_o$ is the frequency centroid of the nonuniform power spectral density plot of the vibration signal and can be calculated from Equation (3).

$$N_o = \left[\frac{\int_o^{f_{max}} f^2\, \Phi_{\dot{Z}\dot{Z}}(f)\,df}{\sigma^2}\right]^{1/2} \qquad (3)$$

When $N(\xi)$ is plotted as a function of $\xi$ with given $\sigma$ and $N_o$ values, a familiar bell-shaped curve results (Figure 4). The height of the bell depends upon $N_o$, and the width of the bell depends upon the value of $\sigma$ (e.g., at $N(\xi) = 0.606\left(\frac{N_o}{2\pi}\right)$, the half-width of the bell measured from the ordinate equals $\sigma$).

Since the parameters $\sigma$ and $N_o$ that determine the character of the composite exceedance curve are actual physical constants of a given vibration power spectral density distribution, one-to-one correspondences exist between the height and the frequency centroid, the

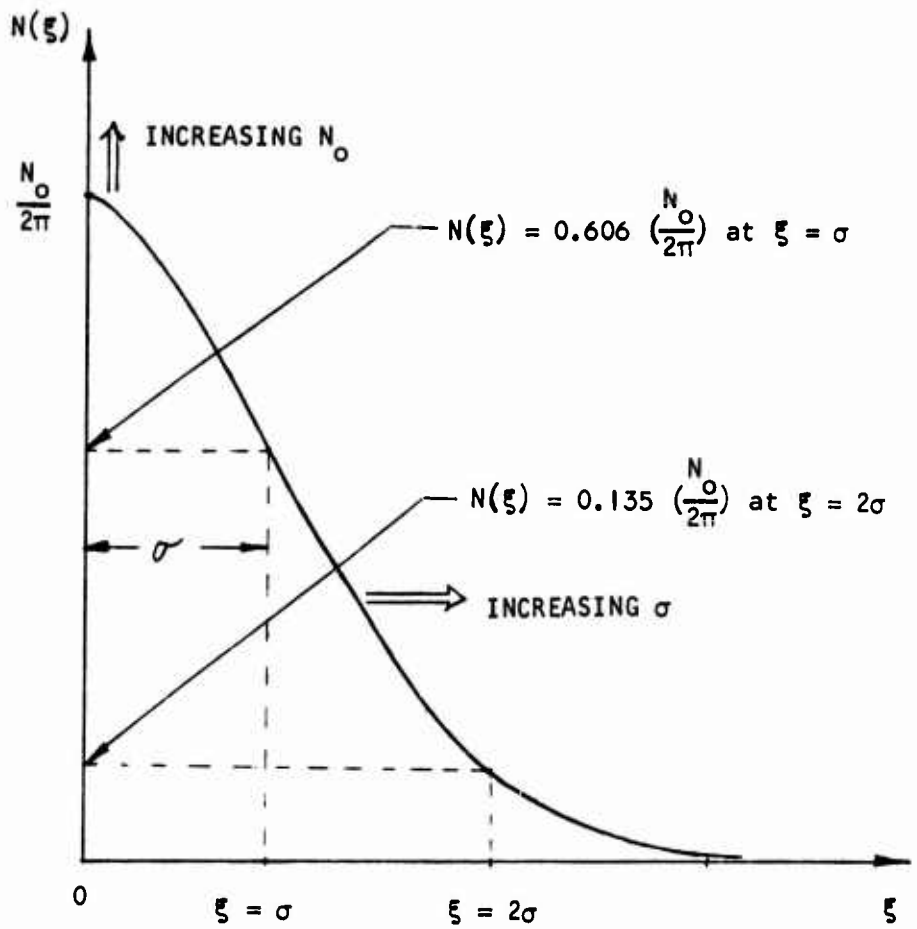



Figure 4. Characteristics of a Composite Exceedance Curve.

width and the total power content, and the shape of the curve and the narrow-band Gaussian property of the signal, respectively. For a normal component, there will be a bell-shaped curve of a given height and width. When a failure condition is developing, either the amplitude of the existing vibration frequencies will be increased, new frequencies outside the narrow band will be created, or both these events will occur. The increase in amplitude of existing vibration frequencies will result in larger $\sigma$ and in turn will increase the width of the curve (or move the curve to the right). The generation of new frequencies will result in higher $N_o$ counts and a deviation from the narrow-band nature of the normal signal. These, in turn, will cause the curve to move upward and also distort the curve from a single bell shape to some more complicated pattern. Therefore, by comparing the actual composite exceedance curve obtained at a certain time instant in operation with a reference exceedance curve obtained from a normal component of the same kind, an abnormal mechanical condition can be detected by examination of the deviations of the actual curve from the reference curve.

Special Application Considerations

An implicit assumption made in the previous discussion is that both $N_o$ and $\sigma$ are constants. In an actual application, the vibration signal is more or less nonstationary, $N_o$ and $\sigma$ being constants only for a short period of time. Additionally, both values also depend upon the operating rpm of the component. Therefore, in order to apply the aforementioned theoretical approach to actual mechanical component status monitoring, and ultimately, prognosis, certain refinements of the method have to be made.

To compensate for the inherent statistical variations of $N_o$ and $\sigma$ among normal units themselves, a sufficient number of normal components must be tested to establish a reliable average reference composite exceedance curve as well as the significant band of variation of $N_o$ and $\sigma$ values. The detection should be made on the basis of comparing the actual curve with an average reference curve and the band of variations of the normal units (Figure 5).

To compensate for the variations due to nonstationariness, the values of $N_o$ and $\sigma$ have to be tagged with a time index, $t_i$, to signify the point of application in the time domain, and only the corresponding reference composite exceedance curve valid at the same time instant shall be used for comparison. This requirement calls for the following actions to Equations (1), (2), and (3).

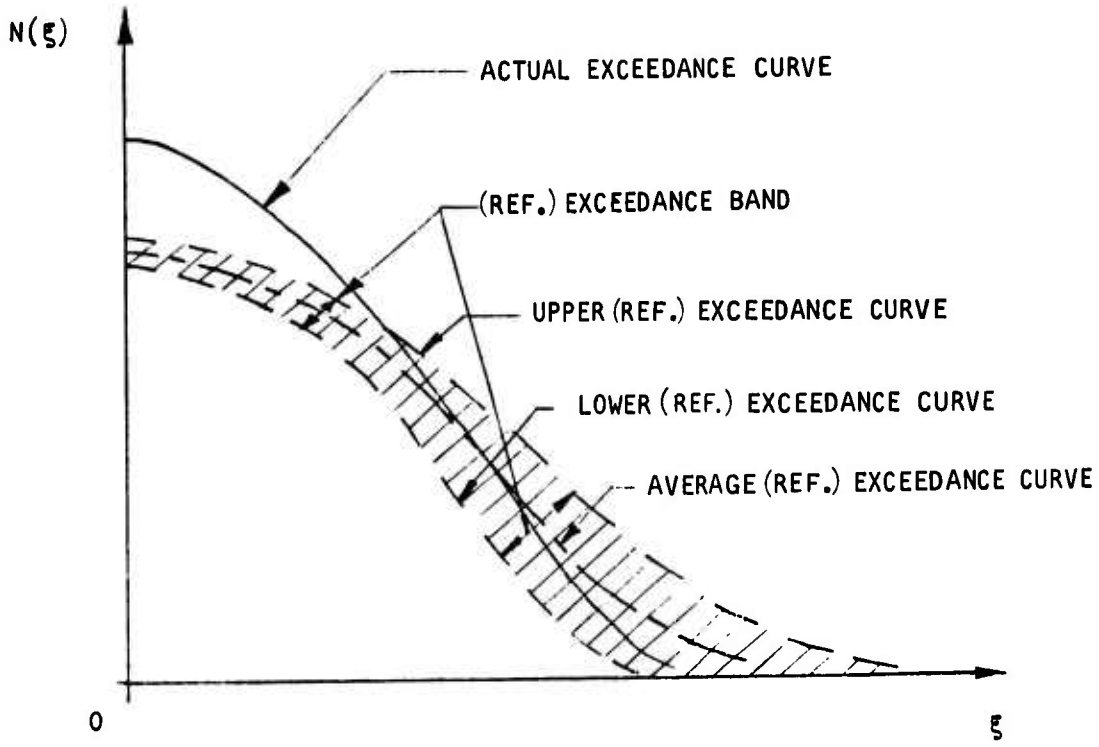



Figure 5. Comparison of an Actual Exceedance Curve With a Reference Exceedance Band.

$$\Phi_{\ddot{z}\ddot{z}}(f) \longleftarrow \Phi_{\ddot{z}\ddot{z}}(f, t_i)$$

$$\sigma \longleftarrow \sigma(t_i)$$

$$N_o \longleftarrow N_o(t_i)$$

$$N(\xi) \longleftarrow N(\xi, t_i)$$

where " $\longleftarrow$ " means "being replaced by". The above piece-wise stationary nature of the vibration signal requires that either a set of reference curves predetermined at some proper time instants $t_i$, $t_{i+1}$, $t_{i+2}$... should be used, or the initial reference curve and band should be updated with time.

To allow proper accounting for testing under various speed conditions, the reference curves should also be properly weighted by the anticipated relative utilization factor of a given speed level (or $\Delta$rpm band) in terms of the ratio of dwelling duration at that level (or in that band) with respect to the total testing time at $t_i$.

Establishment of Reference Composite Exceedance Curves

Without taking into account the statistical variations and speed effects, the basic reference exceedance curves can be established in two ways. They can be generated analytically according to the theoretical equation of $N(\xi)$ with the proper values of $N_o(t_i)$ and $\sigma(t_i)$ employed, or they can be generated by direct counting of amplitude peaks of the vibration waveforms.

Let the duration of the measurement at speed level $\Omega_j$ be $T_j$, and let the number of peaks that exceed or equal a vibration level $\xi$ during this time duration and from the $k^{th}$ test unit be represented by $N_{jk}(\xi, t_i)$. The first approach will require the measurement of $\Phi_{\ddot{z}\ddot{z}}(f, t_i)$ of a normal component at time $t_i$; the desired exceedance curves are calculated according to the following equation:

$$N_{jk}(\xi, t_i) = \frac{N_o(t_i) \cdot T_j}{2\pi} e^{-\frac{\xi^2}{2\sigma^2(t_i)}} \qquad (4)$$

where
$$\sigma^2(t_i) = \int_0^{f_{max}} \Phi_{\ddot{z}\ddot{z}}(f, t_i)\, df \tag{5}$$

and
$$N_o(t_i) = \left[ \frac{\int_0^{f_{max}} f^2\, \Phi_{\ddot{z}\ddot{z}}(f, t_i)\, df}{\sigma^2(t_i)} \right]^{1/2} \tag{6}$$

The second approach will require the division of the amplitude level ($\xi$) into discrete bands (see Figure 6) and then the tabulation of the number of peaks within a band and beyond a band in a manner as illustrated in Table I. The peaks can be identified by the zero slope or by the change from a positive slope to a negative slope. The relationship between the cumulative number of peaks $\geq |\xi|$ and the amplitude bands represents the curve $N_{jk}(\xi, t_i)$ and should be identical with the curve calculated from Equation (4).

To satisfy the additional requirements for application of the method to actual status monitoring, the averaging of the curves from a number of normal units and the weighting of each basic reference curve of different $\Omega_j$ by the corresponding utilization factor will be used. In terms of the basic exceedance reference curve $N_{jk}(\xi, t_i)$ of the $k^{th}$ test unit at speed level $\Omega_j$, the averaging process can be accomplished by the following operation:

$$N_j(\xi, t_i) = \frac{1}{n} \sum_{k=1}^{n} N_{jk}(\xi, t_i) \tag{7}$$

where n is the total number of units under test. Let $T_i$ be the duration of the test run at time instant $t_i$. The utilization factor of speed level $\Omega j$ in the test run at time instant $t_i$ is simply

$$P_j = \frac{T_j}{T_i} \tag{8}$$

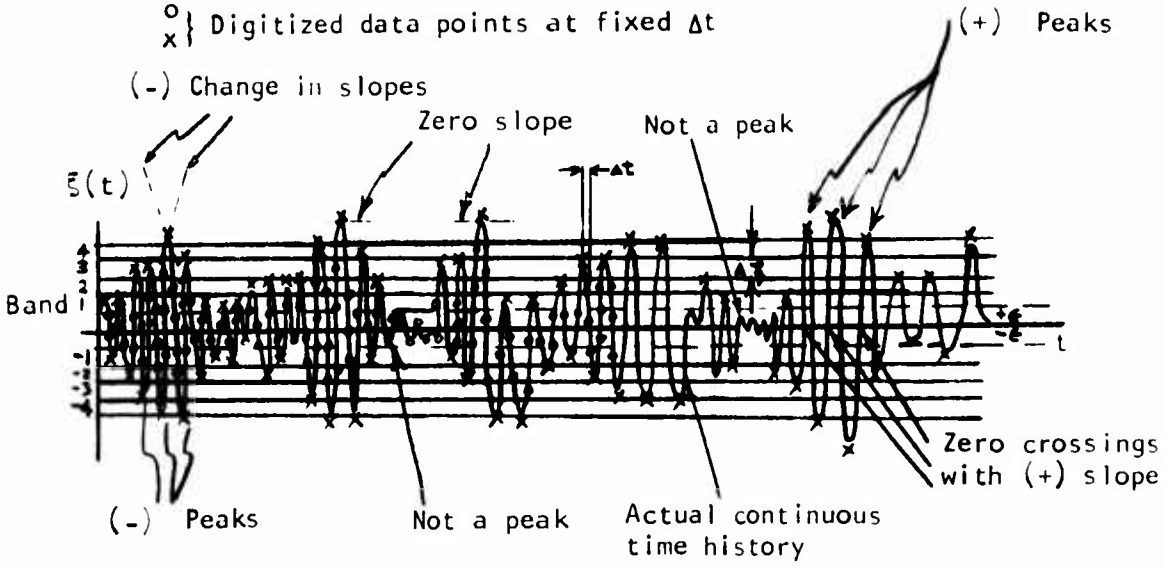



Figure 6. Example of Peak-Counting Procedures.

TABLE I. PEAK-COUNTING RESULTS

| Band Number | ξ | No. of Peaks Within the Band | Cum. No. of Peaks ≥ ξ | Cum. No. of Peaks ≥ \|ξ\| |
|---|---|---|---|---|
| Above Band +4 | $> 4\Delta\xi + \epsilon$ | 5 | 5 | 6 |
| Band +4 | $4\Delta\xi + \epsilon$ | 11 | 16 | 24 |
| Band +3 | $3\Delta\xi + \epsilon$ | 7 | 23 | 36 |
| Band +2 | $2\Delta\xi + \epsilon$ | 9 | 32 | 53 |
| Band +1 | $\Delta\xi + \epsilon$ | 9 | 41 | 72 |
| Band -1 | $-\Delta\xi - \epsilon$ | 10 | 31 | |
| Band -2 | $-2\Delta\xi - \epsilon$ | 8 | 21 | |
| Band -3 | $-3\Delta\xi - \epsilon$ | 5 | 13 | |
| Band -4 | $-4\Delta\xi - \epsilon$ | 7 | 8 | |
| Below Band -4 | $< -4\Delta\xi - \epsilon$ | 1 | 1 | |

Then the weighting of the average reference exceedance curve becomes

$$N(\xi, t_i) = \sum_{j=1}^{m} \left(\frac{T_j}{T_i}\right) N_j(\xi, t_i)$$

$$= \frac{1}{n} \sum_{j=1}^{m} \left(\frac{T_j}{T_i}\right) \sum_{k=1}^{n} N_{jk}(\xi, t_i) \qquad (9)$$

Equation (9) represents the final form of the averaged and weighted reference composite exceedance curve that should be used for comparison with the measured exceedance curve at a given time instant $t_i$ in operation in order to determine the status or status degradation of a mechanical component.

Generation of Actual Composite Exceedance Curves During Operation

Since the condition of a mechanical component can be expected to degrade gradually during the time of its use, an increase in vibration amplitude and/or the generation of new frequencies can generally be anticipated as the component usage time accumulates, even though the overall health status of the component may very well be considered as normal. For this reason, the narrow-band Gaussian assumption of the vibration data will no longer be true as time goes on. The theoretical approach of generating the exceedance curve by means of computation according to Equation (1) will, therefore, be invalid for the actual exceedance curve. This leaves only the possibility of generating the actual status-indicative exceedance curve during operation in the form of the second approach, i.e., direct counting of amplitude peaks of the measured vibration waveforms.

The peak-counting technique involves the performance of a number of routine steps such as sampling, digitization, quantizing the amplitude bands, measuring the slope, comparing the slope, and accounting for the peaks over a large number of data points. These tasks may be difficult to perform manually but involve no problem for electronic devices. With the speed capability of present-day digital electronics, the peak-counting technique needed to generate the actual exceedance curve can be performed in real time by a small, general-purpose digital computer or a special-purpose discrete logic system.

Depending upon the accuracy, the speed levels involved in a measurement run, and the duration of each measurement run, the actual exceedance curve counted during operation may or may not need the weighting with the speed utilization factors. Let the unweighted actual exceedance count function measured during the operation be represented

by $C_j(\xi, t_i)$, where j designates the speed level $\Omega_j$. The weighted count function will then become

$$C(\xi, t_i) = \sum_{j=1}^{\hat{m}} \frac{\hat{T}_j}{\hat{T}_i} C_j(\xi, t_i) \qquad (10)$$

where $\hat{T}_i$ = duration of the measurement run at time instant $t_i$

$\hat{m}_i$ = number of speed levels inolved in the ith measurement

$\hat{T}_j$ = dwelling time at the $j^{th}$ speed level during the ith measurement

By comparing either $C_j(\xi, t_i)$ or $C(\xi, t_i)$ against $N(\xi, t_i)$ and the associated upper and lower bounds, the health condition of a mechanical component and its progressive change can then be concluded.

Acoustic Emission Signature Analyses

Acoustic emission is a generic term that represents all the airborne sound pulses and structural-borne elastic waves in a solid material caused by a deformation process in the material. This term includes both the audible sonic signals and the nonaudible ultrasonic signals thus generated.

The audible sounds are commonly known as "tin cry," which is produced mainly by the twinning deformation in the tetragonal and hexagonal crystallographic structures. The nonaudible ultrasonic sounds are mostly created by the sudden reorientation of large grains in a polycrystalline material and certain phase transformations such as are associated with a heat treatment.

The ultrasonic sounds that are generated by the sudden grain reorientation are the most important type of signals from a failure monitoring and detection point of view because of their ubiquitous existence, their significant signal intensity, and their ability to signify the early failures such as microscopic crystalline dislocations, slips, and the breakage of microscopic fibers. Acoustic emission is an irreversible process. It is a phenomenon that can be generated anywhere within the volume and on the surface of the testing material. It thus can signify both the surface and subsurface flaws.

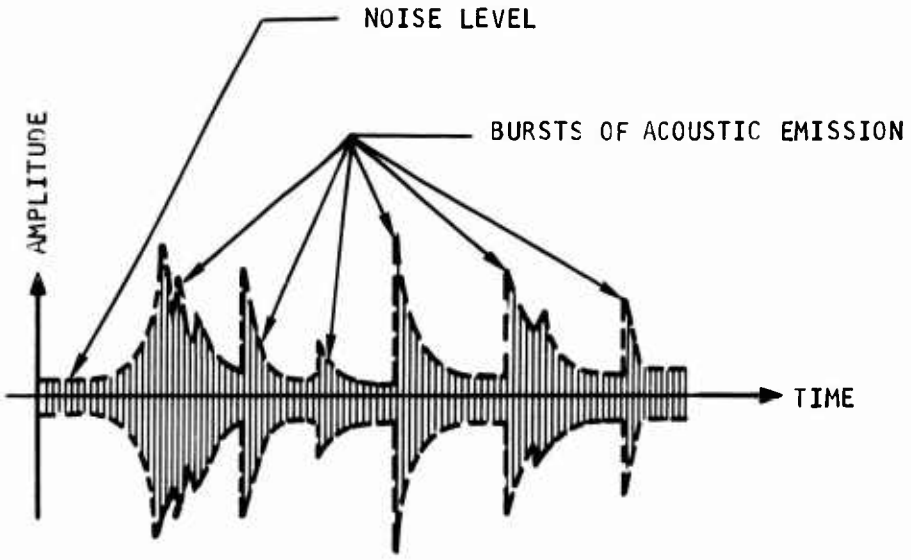



Figure 7. Typical Acoustic Emission Signals.

So far, two types of acoustic emission signals have been observed: burst type signals and the continuous emission type signals. The burst type signals are of high frequency and short duration. Figure 7 shows a typical waveform of this type of signal, which is composed of repetitive bursts of decaying high-frequency oscillations on top of noise amplitudes.

The amplitude of the acoustic emission signals is a function of the strain rate. The emission rate is found to be correlatable to the number of fine slip lines or the crack growth rate existing in the material in many cases; however, the exact definition of emission rate is still an area that needs unification. In some applications, this is defined as the number of signal bursts above the noise level per unit time. In other applications, it is defined as the number of decaying oscillations per unit time.

The difference between these two definitions of emission rate is that the former treats all the bursts of different amplitudes on an equal basis, while the latter weights the emission rate by the severity of the flaw because more cycles of oscillation will be needed to ring down to the same threshold level. The total emission, which is the time integral of the emission rate, has been found to be indicative of the percentage of the cracked fiber or the crack length (Reference 4). Equation (11) shows a theoretical relationship between the total emission (N) and the percentage of cracked fiber ($d\omega/d\epsilon$):

$$N = B \int_{\epsilon_o}^{\epsilon} \left(\frac{d\omega}{d\epsilon}\right) \ln\left(\frac{\epsilon}{\epsilon_o}\right) d\epsilon \qquad (11)$$

where $B$ = a proportionality constant

$\epsilon$ = the strain level

$\epsilon_o$ = the threshold strain level

The continuous emission type signals are of low amplitudes. It is believed that the continuous emission signals are created by the stress-induced dislocation movement constantly going on within the material. Because of the low amplitude, they are easily contaminated by the noise and become very difficult to be transduced. Consequently, the use of this type of signal is limited even though it has a greater potential for monitoring and for prognostic applications.

Because of the difficulty encountered in the sensing of the continuous type of acoustic emission signals, current analysis of the acoustic emission signals is limited mainly to the burst emission type.

Figure 8 shows a typical arrangement used in the analysis setup. This includes a piezoelectric type sensor to transduce the structure-borne elastic waves into electrical charges, a charge conversion preamplifier to

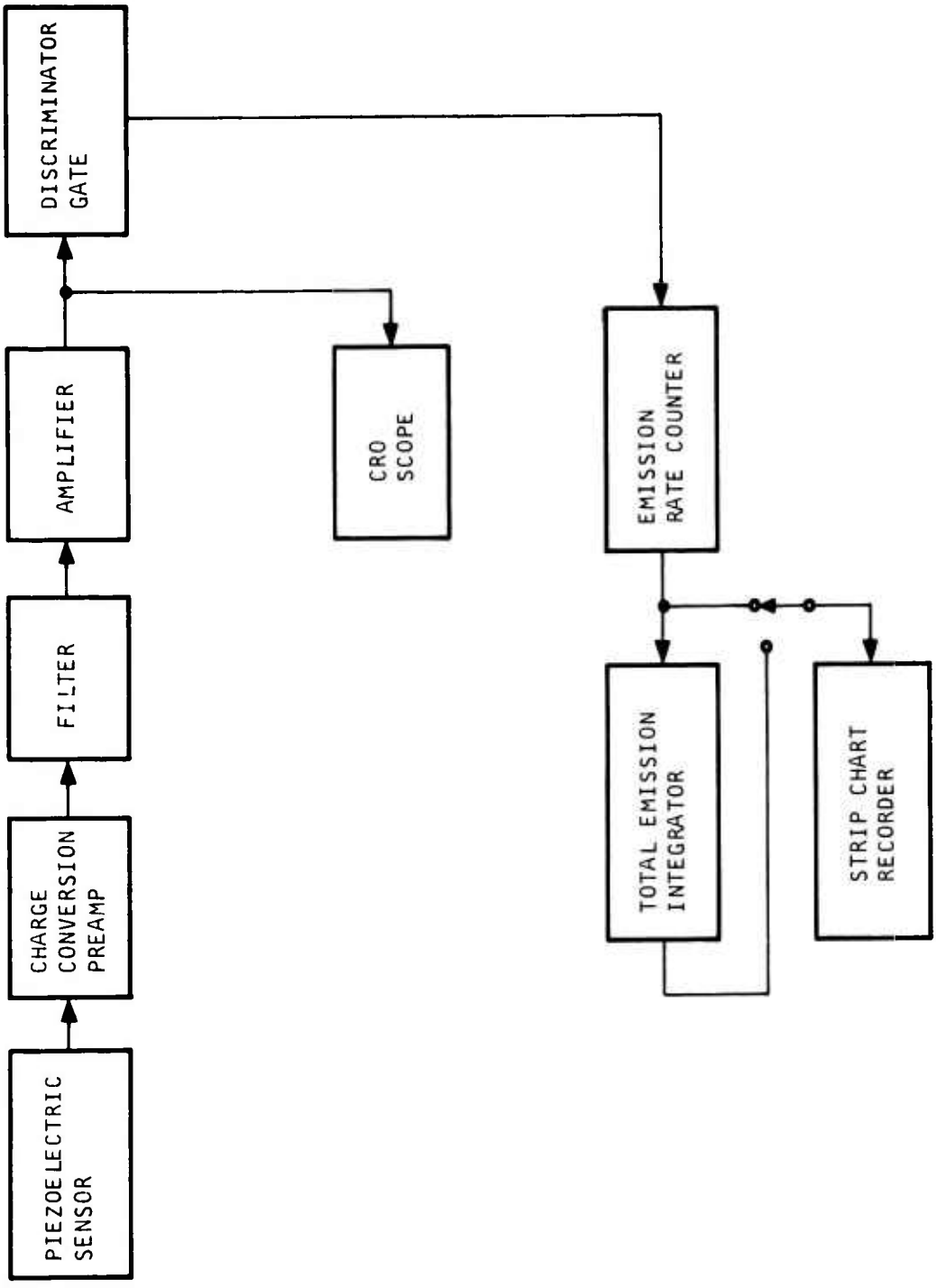



Figure 8. Typical Acoustic Emission Analysis Setup.

change the charge variations into voltage variations, a high-pass or band-pass filter to remove the unwanted noises, an additional amplifier to bring up the signal amplitude, a discriminator gate to shape the incoming acoustic emission signals into pulses according to the definition of the emission rate, a CRO scope to allow the viewing of the waveform, an emission rate counter to count the number of occurrences per unit time of the pulses coming out from the discriminator gate, a total emission integrator that sums up the grand total counts of the variable emission rate, and a strip chart recorder that permits the graphical representation of the emission rate vs time or the total emission vs time.

To measure acoustic emission accurately up to 100 kHz, a sensor resonant frequency of 300 to 500 kHz is necessary. Existing accelerometers are designed tc have a resonant frequency up to 150 kHz. Consequently, measurements of acoustic emission with frequency components higher than 30 kHz become the weighted results of the signal by the sensor frequency response. Total amplifier gain needed in the equipment is around 10,000. The filter in general is a tunable one so that the band of frequencies desirable can be selective. The bandwidth of the filter is also variable depending upon the application; it ranges from 1 or 2 kHz to about 40 kHz. The attenuation required is in the area of 50 db per octave.

The discriminator gate compares the amplitude of the amplified acoustic emission signals against a preset threshold level, which represents the noise and then shapes those of exceedance into better defined pulses for emission rate counting. In general, the emission rate is expressed in terms of the counts of exceedance pulse per second; however, in testings under a cyclic load, emission rate sometimes is also expressed in terms of the number of counts of exceedance pulses per cycle of the load application. So far, only the emission rate and the total emission have been used to assess the flaw condition. They are plotted through a strip-chart recorder in the form of a discrete function of time.

Although the possibility of determining the fault location by use of seismic triangulation technique and multiple sensor deployment has been suggested by researchers in this area, no existing literature can be found to provide substantiation. However, this could endow the diagnostic power to the acoustic emission analysis technique if such a triangulation determination can be developed.

Although there are many possible applications, all dependent upon user's innovation, the following are some typical major applications of the acoustic emission signature analysis failure monitoring technique.

Acoustic Emission Analysis for Cyclic Fatigue Failure Monitoring

Initial tests with prepared specimen have been performed in a laboratory environment. High-cycle or low-cycle tension or tension-compression loads were applied to the prepared specimen. Total emission and emission rate were plotted against the load cycles. Figure 9 shows a typical plot from such a test. The measured crack growth rate and the crack length were also plotted on the same sheet to show the correlations.

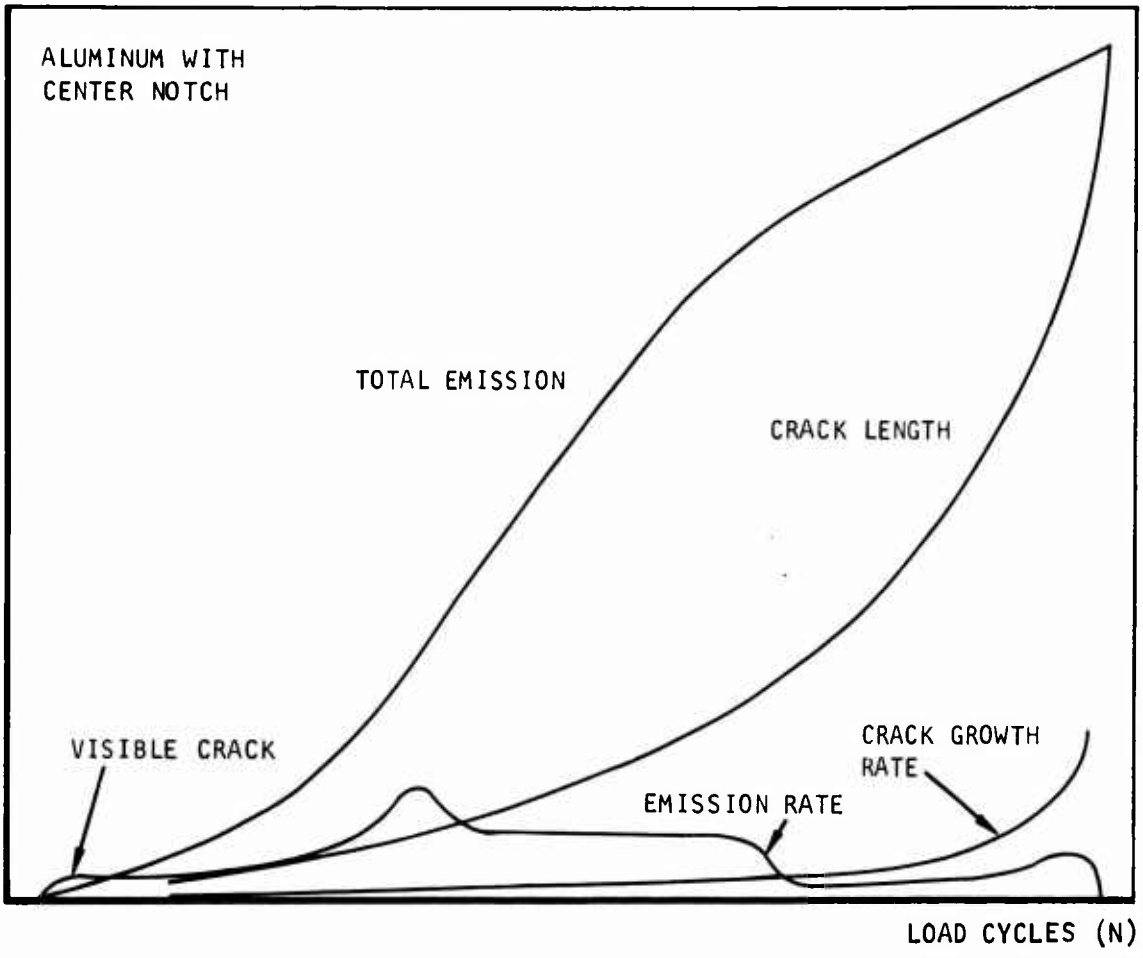



Figure 9. Mechanical and Acoustic Emission Fatigue Test Results.

The existence of a consistent exponential relationship between emission counts and the stress intensity allows the possibility of describing a through crack in terms of acoustic emission process. It is hoped in this area of application that, by testing a limited number of test specimens, the prognostic predictions of the useful life remaining can be inferred through statistical means.

Acoustic Emission Analysis for Monitoring Pressure Vessel Fracture

Deformation of a pressure vessel under noncyclic tension stress has been monitored by the acoustic emission technique (Reference 6). Figure 10 shows a typical acoustic emission plot. When the pressure level in the vessel reached a certain level, a sharp increase in emission rate was observed. After passing the yielding point, the emission rate began to slow down. This emission rate will increase rapidly again just prior to the complete fracture. The two rapid increases in emission rate thus provided an early warning capability of the occurrence of yield and the complete failure.

Acoustic Emission for Stress Corrosion Monitoring

When a stressed mechanical component undergoes a chemical corrosion process, acoustic emission bursts are also being created. Tests performed on laboratory specimens (Figure 11) showed that there exists an initial warning of the beginning of the corrosion process as signified by the sharp emission rate. The full development of a crack is indicated by the high counts obtained after the initial sharp rise. The acoustic emission rate then begins to taper off after passing the second high count.

Acoustic Emission for Monitoring Welding Flaws

Acoustic emission also can be used to detect the formation of cracks during the welding process or during the post-welding cooling period. A simulation test was performed using tantalum and titanium impurities to purposely introduce the flaws during welding. Figure 12 shows the result of such a simulation test.

The dramatic difference in acoustic emission from the faulty welds becomes apparent as indicated in the diagram. Since the weld cracking generally produces a rather high-amplitude long-duration emission signal, it becomes possible to suppress the short-duration arcing noise by electronic filters.

Acoustic Emission for Monitoring Bearing Flaws

In ball bearings, both types of emission have been observed (Reference 8). The burst type emission occurred when flaws were initiated in the inner race, the outer race, and on a ball. The so-called continuous signal has been observed in a dry bearing operated at high speed and in an oscillating bearing operated through a partial circle at low speed with the application of heat in a vacuum under stress. In the

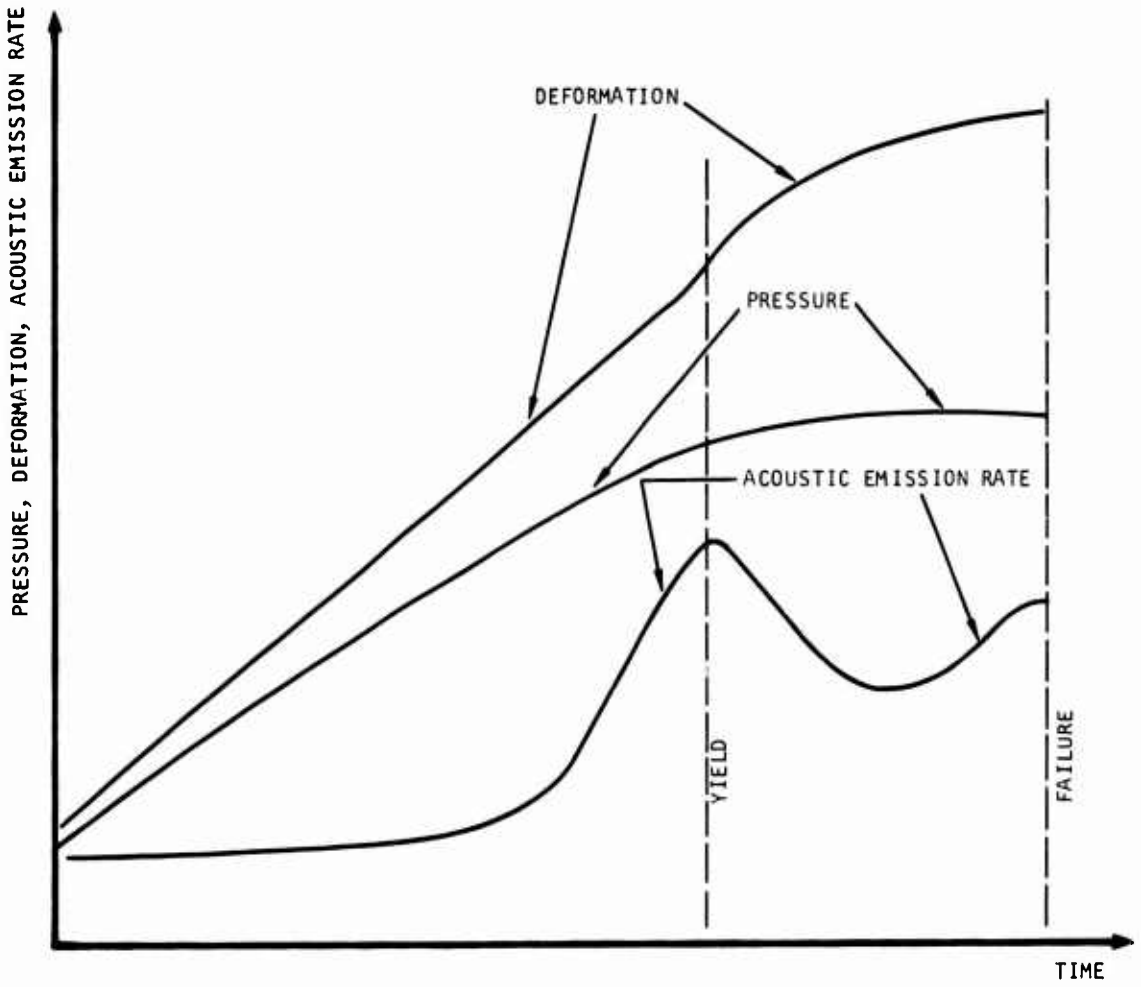



Figure 10. Typical Pressure Vessel Test Acoustic Emission Plot.

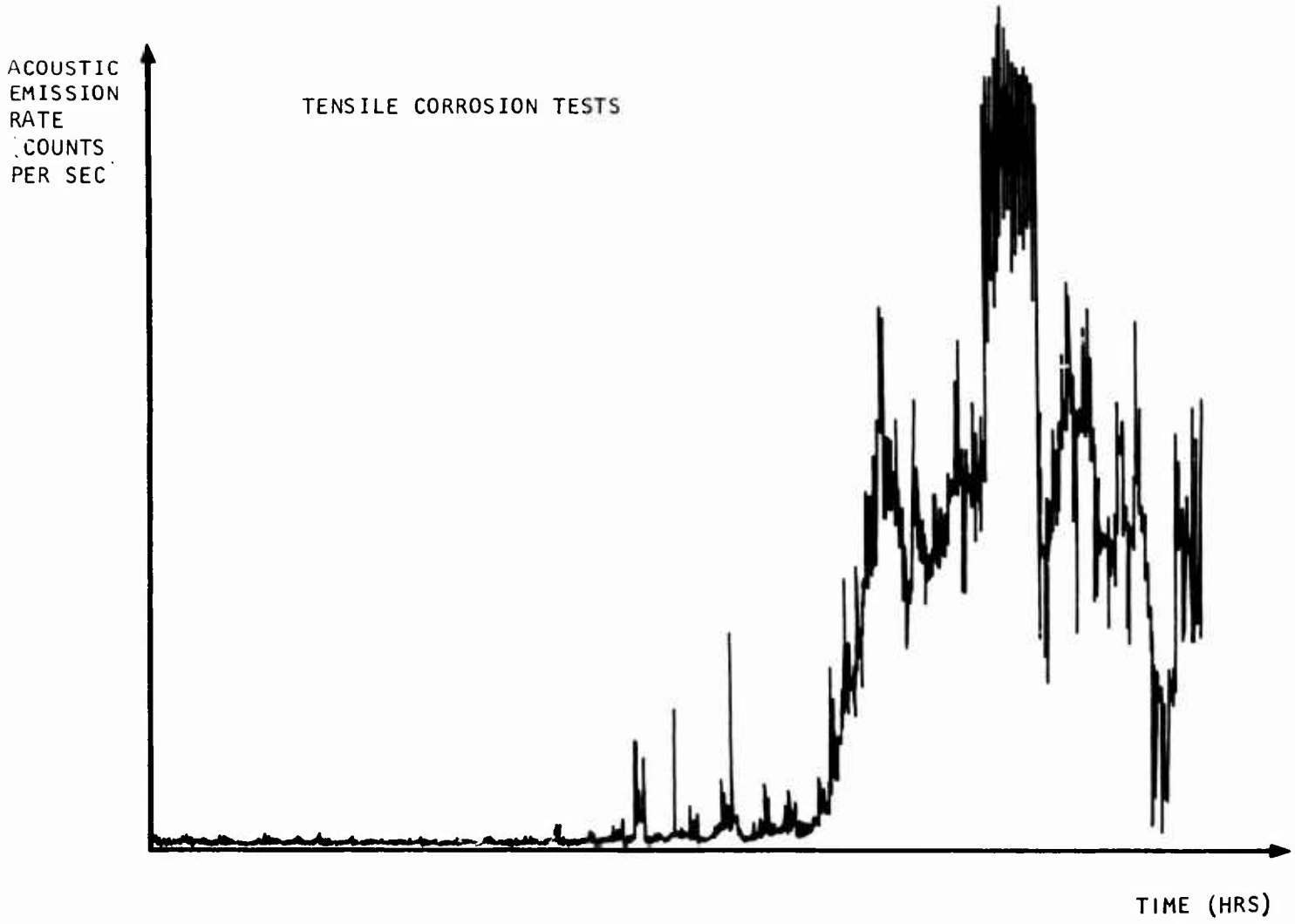
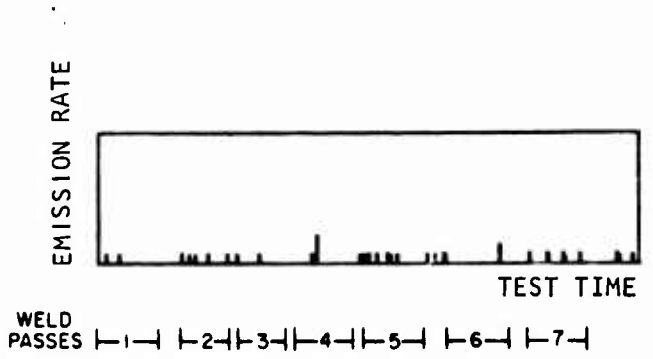
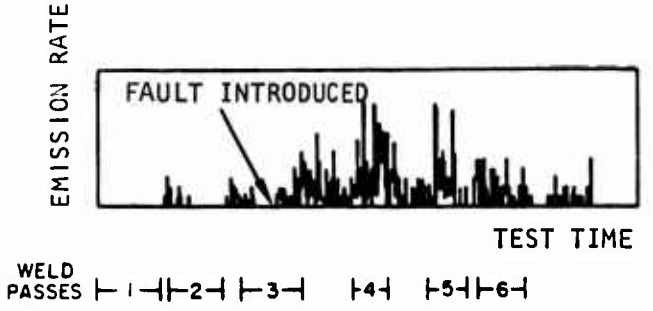


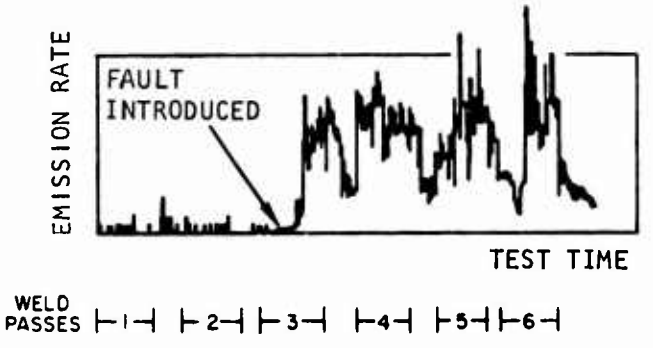

Figure 11. Sample Result of Stress Corrosion Tests.




a. NO FAULT INTRODUCED




b. LIGHT FAULT INTRODUCED




c. HEAVY FAULT INTRODUCED

Figure 12. Acoustic Emission Generated by Weld Faults.

latter bearing test, wherein one bearing was failed using only dry lubricant, a very high rate of acoustic emission signals occurred, during and prior to actual seizure of the bearing. A great deal of wear occurred in this particular bearing due to oxidation of bearing surfaces and subsequent scaling of the oxidized layer. A large number of burst signals, as many as 30,000 per minute, were observed during this wear process.

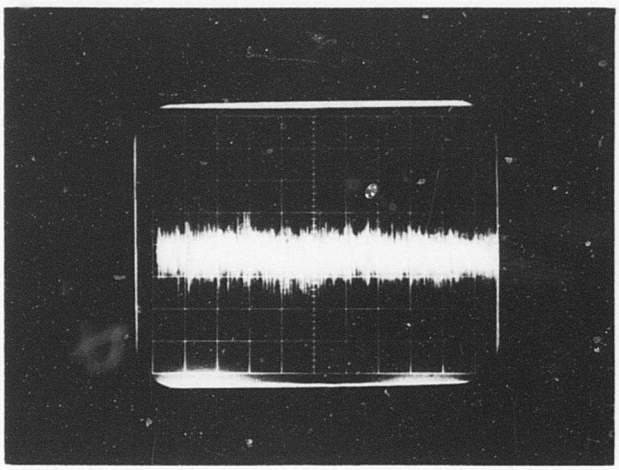
However, not all the burst type emissions in a bearing are caused by acoustic emission; large numbers of them are caused by the rotation-initiated impact shock pulses and damped resonant frequencies of the bearing components. Figures 13 and 14 show the comparison of waveforms of a possible mixture of the acoustic emission signals and shock pulses before and after minute incipient damage was done to the No. 1 main bearing of a J57 turbojet engine used as a test vehicle for the USAF turbine engine diagnostic program conducted by AiResearch (Reference 2).

Wear Particle Generation Monitoring

From a prognostic viewpoint, comparison of the merit or the information content of different wear particle generation monitoring methods is based on which method provides information most closely related to abnormal processes, i.e., processes preceding failure or life expenditure in a transmission. The argument that the spectrometric oil analysis procedure (SOAP) used exclusively or in conjunction with other monitoring techniques leads to an excessive number of false alarms, while a significant number of failures remains undetected, may be explained by the particle-size range selective character of the spectrometric oil analysis method. Its detection capability is concentrated in the particle size range more typical of wear conditions than of fatigue conditions, which figure heavily in transmission component failure modes. In contamination studies of hydraulic and fuel systems, correlations between particle size and the process types in which these particles are generated have been defined. Such a summary recently has been given by Beerbower (Reference 9), who also attributed the upper particle size range limitation (10 micron) of SOAP to particle settling time as a function of the time elapsed between system shutoff and sampling.

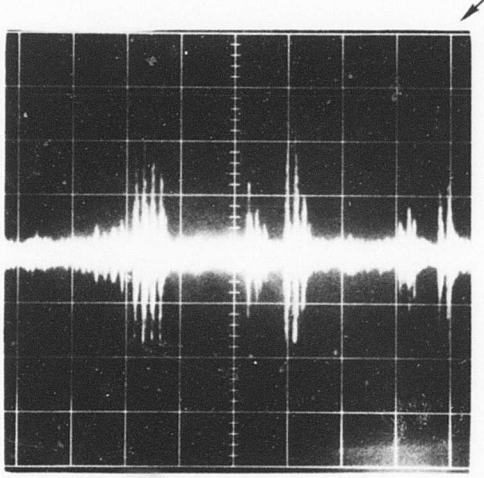
As these contamination studies have shown, typical particle sizes vary between molecular, as generated by corrosion processes, to over 100 microns for rolling contact fatigue, with adhesive wear and sliding fatigue processes assuming a median position characterized by particle sizes in the order of from 1 to 5 microns. For scuffing and abrasive wear processes, particle sizes appear to follow a log-normal distribution as defined by Cole and Fitch (References 10 and 11).

Mathematical modelling methods can be used for the quantitative systemization of wear particle generation processes. Examples are described by Cole and Fitch (References 10 and 11) and Beerbower (References 12 and 13). For wear processes of different types such as corrosive, fatigue, adhesive, and abrasive, Beerbower found that the individual



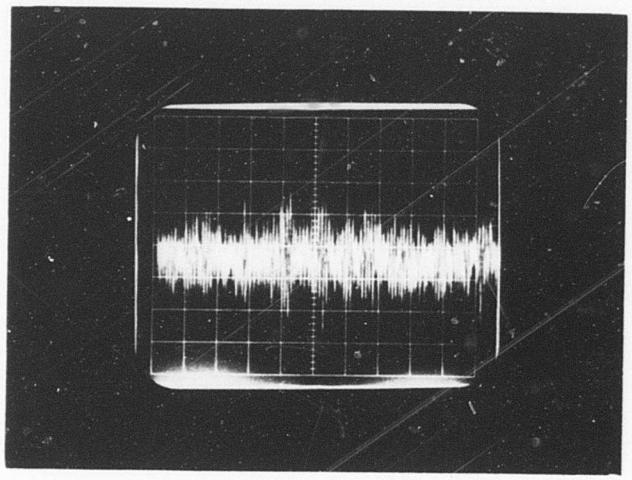
WITHOUT HAIRLINE SCRATCH ON NO. 1 BEARING INNER RACE,
0.2 V/VERTICAL DIVISION, 2 MSEC/HORIZONTAL DIVISION

SENSOR OUTPUT
WAVEFORMS



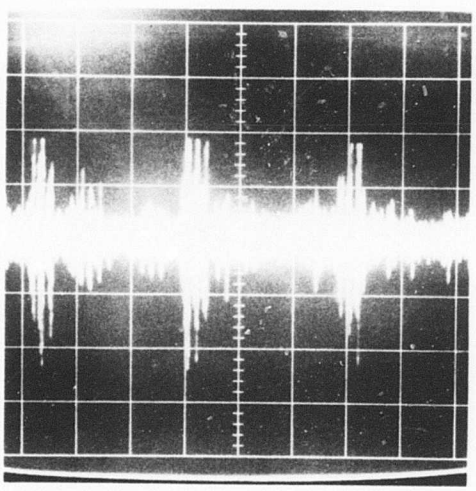
WITH HAIRLINE SCRATCH ON NO. 1 BEARING INNER RACE,
1 V/VERTICAL DIVISION, 10 MSEC/HORIZONTAL DIVISION

Figure 13. Detection of Incipient Bearing Failure by Acoustic Emission/Shock Pulse From a High-Frequency Accelerometer Output.



WITHOUT HAIRLINE SCRATCH ON NO. 1 BEARING INNER RACE,
0.2V/VERTICAL DIVISION, 2 MSEC/HORIZONTAL DIVISION

SENSOR OUTPUT
WAVEFORMS



WITH HAIRLINE SCRATCH ON NO. 1 BEARING INNER RACE,
1 V/VERTICAL DIVISION, 10 MSEC/HORIZONTAL DIVISION

Figure 14. Detection of Incipient Bearing Failure by Acoustic Emission/Shock Pulse From a Lightly Damped Sonic Accelerometer Output.

mathematical models describing each type of wear process can be reduced to relationships of the following common format:

$$V/d = Kd^m \times W^n \times U^p \qquad (12)$$

where V = wear volume
d = distance of wearing components traveled since test initiation
W = load
U = velocity of wearing components

As determined by experimentation, the exponents m, n, and p are different for each of the individual wear processes considered. A listing of these exponents is given in Table II.

TABLE II. Wear Process Equation Exponent Values

| Wear Mode | Exponents | | |
|---|---|---|---|
| | m | n | p |
| Corrosive | 0 | 1.7/2.7 | -0.9/-0.5 |
| Fatigue | -0.53 | 2.12 | 0 |
| Adhesive | 0 | 1.93 | -0.60 |
| Abrasive | -0.33/0 | 1 | 0 |

Since both corrosive and adhesive wear processes are characterized by m = 0, n > 1.7, and not-too-different p values, the difficulty of distinguishing between these two wear modes would exist. On the other hand, a fatigue process would be relatively easily identifiable.

Since most wear processes are of the composite-cause type, Beerbower proposed to blend the modes on a nonlinear basis, as expressed by the compound equation

$$(V/d) = (V_f + V_a + V_c + V_{fa} + V_{fc} + V_{ac} + V_{fac})/d + V_d/d \qquad (13)$$

where subscripts f, a, c, and d refer to fatigue, adhesion, corrosion, and abrasion, respectively.

In view of helicopter transmission design trends toward compartmentalization of individual major subassemblies, with possibly a particle size-discriminating electric chip detector installed in each compartment, relationships of a type similar to those expressed by Equations 12 and 13, suitably adapted to the wear mechanisms in aircraft transmissions, could be useful for life prognosis. On the basis of wear rate, they may allow determination of the type of wear process that is dominant in the mechanical degradation process of a specific transmission system component. Particle size-discriminating instrumentation (such as advanced types of electric chip detectors or conductive filters with graduated mesh size)

could provide supplementary inputs to mathematical models designed to represent wear processes of different types.

Present efforts in the area of rationalizing fluid-flow particle contamination conditions, primarily with the objective of arriving at meaningful, quantitative criteria for contamination condition tolerance units for hydraulic systems, are of interest in view of study objectives since the concepts developed in these efforts correlating particle concentration and size distribution could be used to characterize a changing lubricant contamination condition in a gearbox or transmission in a more comprehensive manner than in current practice.

Figure 15 shows the application of statistical methods to cumulative particle count data (particularly in the form of the count linearization model), plotting on the ordinate the number of particle sizes greater than indicated on the abscissa in the form of a log-log$^2$ graph. The cleanliness level nomograph shown in Figure 16 could be used in a suitably adapted form to follow closely a mechanical system degradation process for prognostic purposes.

The nomograph shown in Figure 16 uses the number of particles greater than 10 and 20 microns to define the distribution curve, which is represented by a point on the chart. This distribution point defines (1) a cumulative particle size distribution curve, (2) an average volume for all particles contained in the distribution, and (3) the fluid gravimetric level.

The curves on the nomograph represent constant gravimetric levels; the straight lines of slope one, the constant average particle volumes. The four areas of the nomograph ("ideal" or indefinite life, "marginal" or life factor limitating, "poor filter" or unsatisfactory life, and "change oil" or high content of fine particles) can be determined experimentally, for instance by endurance testing.

The methods presently available for lubricant-contained prognostic signature analysis covering both lubricant particle content and physical/chemical conditions are listed in overview form in Table III. Of the fourteen methods shown, only three (items 9, 10, and 11) are implementable by operationally suitable sensors. Millipore filtration and filter inspection methods may be implementable on the flight line. For the nuclear fluorescence and ferrographic methods, operational sensors are under development.

From a prognostic viewpoint, ferrography, microscopy, and millipore filtration have been assigned a highest merit rating since these methods have a most comprehensive coverage of a given lubricant particle contamination pattern in terms of concentration, particle size range, and particle size distribution. Materials identification or determination of materials-related concentration, however, is not possible using these methods. This complementary information would have to be obtained from either emission spectrography or the nuclear fluorescence method.

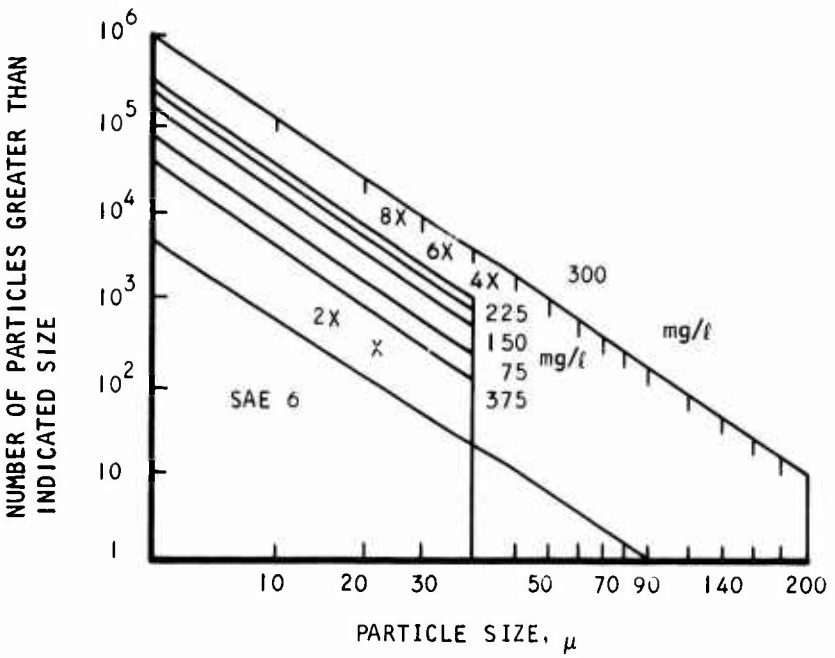



Figure 15. Contaminant Concentrations for Gravimetric Tolerance Tests. (Figure 5 of Reference 11)

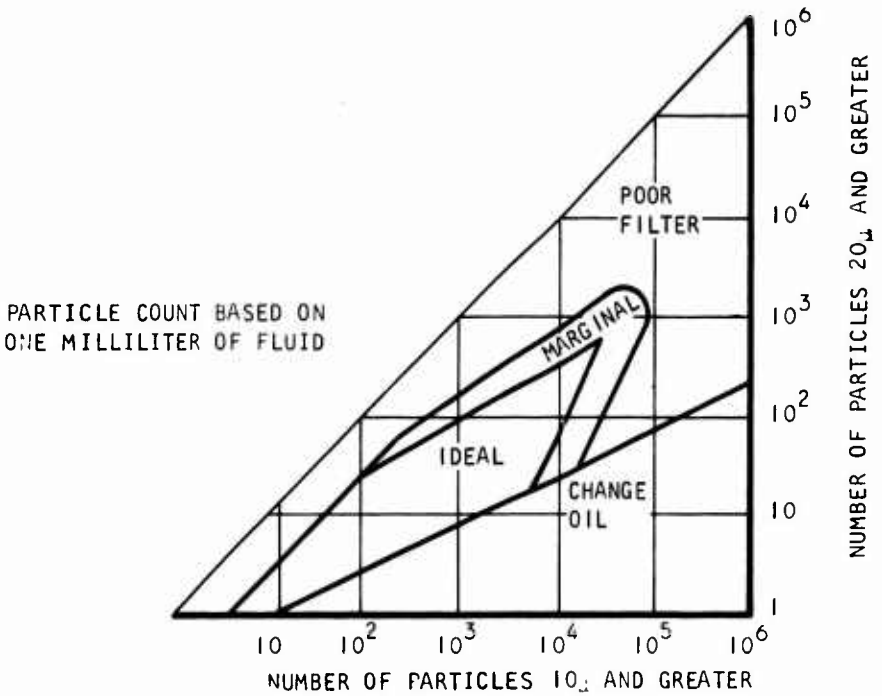



Figure 16. Fluid Cleanliness Level Requirements.

## TABLE III. LUBRICANT-CONTAINED PROGNOSTIC SIGNATURE ANALYSIS METHODS

| Item No. | | Analysis Method | Particle Size (μ) | Primary Data Source | Material | Concentration | Size Class. | Size Distribut. | Size Distribut. Rel. Concentrat. | Physical Characteristics | Prognostic Merit | Op. Implement. Potential |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | Metallic particle content: | | | | | | | | | | |
| 1 | Micro-techniques | Emission spectrography (soap) | <1 | Extract. sample | High | Medium | Medium | None | None | None | Medium | None |
| 2 | | Optical dispersion/attenuation | <10/<40 | Operat. sensor | None | High | Medium | None | None | None | Medium-high | High |
| 3 | | X-ray fluorescence | All | Operat. sensor (devel.) | High | Medium-high | None | None | None | None | Medium-high | High |
| 4 | | Ferrography | <5 | Extract. sample | Mat. class only | High | High | High | High | Medium | High | None |
| 5 | | Magnetoscope | - | Extract. sample | Mat. class only | Medium | Medium | Medium | None | Medium | Medium | None |
| 6 | | Microscopy (bichromatic) | <1 | Extract. sample | Mat. class only | High | High | High | High | High | High | None |
| 7 | Macro-techniques | Millipore filtration | >1 | Extract. sample | None | High | Medium | High | High | Medium-high | High | None |
| 8 | | Filter inspection | All | System filter | High | Medium | High | None | None | High | Medium | None |
| 9 | | Conductive filtration | >10 | Operat. sensor | Mat. class only | Medium | High | None | None | None | Medium | High |
| 10 | | Electr. chip detection (qualitative) | All | Operat. sensor | Mat. class only | None | None | None | None | None | Medium | High |
| 11 | | Electr. chip detection (quantitative) | All | Operat. sensor | Mat. class only | None | Medium | None | None | None | Medium-high | High |
| | | Lubricant physico/chem. condition: * | | | | | | | | | | |
| 12 | | Acidity test | | Extract. sample | | | | | | | Medium | None |
| 13 | | Infrared photospectrometry | | Extract. sample | | | | | | | Medium | None |
| 14 | | Chromatography | | Extract. sample | | | | | | | Low | None |

*Prognostic information content:
Thermal degradation additive depletion
$H_2O$ content, oxidation, additive depletion,
chemical composition

Considering the various analysis methods listed in Table III, as well as the particle content rationalization methods of the type described in References 12 and 13, the question arises if particle size distribution characteristics could not be combined with material-related concentration characteristics in a manner providing a prognostically meaningful pattern. For ferrous materials, the ferrographic method described by Westcott in Reference 14 permits particle-size distribution-related data to be obtained. The ferrographic method, however, appears to be limited to a particle size range not exceeding several microns that would exclude a sector of the overall particle distribution spectrum that could be of particular prognostic significance. Where operational sensor design types capable of giving information correlatable to size distribution are concerned, the conductive filter type of sensor with graduated mesh size might come closest to meeting this requirement. The size-discriminating electrical chip detector also might have a similar potential.

## Lubricant Physical-Chemical Condition

Changes of lubricant physical-chemical condition can be either the result of normal lubricant wear (i.e., changes in lubricant properties due to prolonged use in a system whose components and component interactions have remained in the normally expected, or design, condition during use) or the result of abnormal component condition changes. Since the latter are of specific interest for prognosis, changes of lubricant property patterns from normal could represent valuable prognostic information.

Traditionally, the major lubricant analysis methods suitable for the detection of lubricant physical/chemical condition changes use the following properties: (1) viscosity, (2) total acid number, (3) conductivity, (4) color (or light transmittance), and (5) density.

In addition to these properties, additive status, water content, and oxidation can be determined by infrared spectroscopy, gas chromatography, and X-ray fluorescence. The following is a summary of the capabilities of the more important of these analysis methods in view of observed lubricant property changes in aircraft turbine engine use.

### Viscosity

Determination of viscosity by the standard capillary viscosimeter method specified in ASTM D 445 can be made with a minimum accuracy of 0.23 percent, corresponding to a change of approximately 0.07 centistokes for Type II lubricants. Since lubricant deterioration could be considered negligible at a total change of viscosity of 0.5 centistokes, the capability of this method can be considered adequate. Lubricants used in aircraft turbine engines for 1,000 flight hours have shown viscosity changes of about +2.3 percent. Data typical for helicopter transmissions have not yet been found in the literature. Since the accuracy of the viscosity measurement method can be increased by use of electronic time measurement and control, modification of the

nonconclusive findings of the attempt to correlate lubricant viscosity and transmission failure rate of Reference 15 by such refinements would be of interest.

Total Acid Number

As pointed out by Jantzen (Reference 16), the problems encountered in determining acid number changes are centered around the fact that minute differences in acid number must be detected, whereas the scatter range of the acid number of aged lubricants is considerable. With insufficient definition, the presence of several acids of different acidity leads to an indistinct equivalent point. Recently developed improved analysis methods make it possible to determine changes in TAN of 1 percent above the initial value.

Also as pointed out by Jantzen, acids generated during lubricant aging have a low boiling point. Acid number analysis of the volatile products of aircraft engine lubricants resulted in acid numbers around 200, which means that vaporization effects could be misleading factors in determining lubricant degradations on the basis of total acid number using conventional methods. While this aspect would be considerably more pertinent to the use of lubricants in aircraft turbine engines than in transmissions, it could prove meaningful when the latter have been operating under marginal lubrication conditions at high temperature levels.

Conductivity

With this analysis method, based on the difference in the order of several powers of ten between the extremely low conductivity of lubricant base stocks and polar-structured acids, definitions in the range of

$$+0.1 \times 10^{-10} \left[\Omega^{-1} \times cm^{-1}\right] \qquad (14)$$

can be obtained. Conductivity changes of about $8.5 \times 10^{-10}$ have been observed in 1,000 hr of aircraft engine operation with a mean conductivity change rate of $0.03 \times 10^{-10}$ per 100 hr of operation.

Color or Light Transmittance

With the advent of in-line sensing systems monitoring lubricant color, the capability of laboratory methods in detecting color or transparency changes as an indicator of lubricant degradation is of comparative interest. Using a two-beam spectrophotometer, reproducibility to less than ±1 percent can be obtained. As shown in Figure 17, transparency changes as a result of prolonged lubricant use depend on wavelength used, with larger transparency changes occurring at smaller wavelengths. As illustrated in Figure 18, 1-hr operation of lubricant at 100°C temperature level after a 72-hr use time under normal cruise conditions reduced transparency at 500 μm wavelength to about 30 percent of its original value. Thus it is expected that this method will be

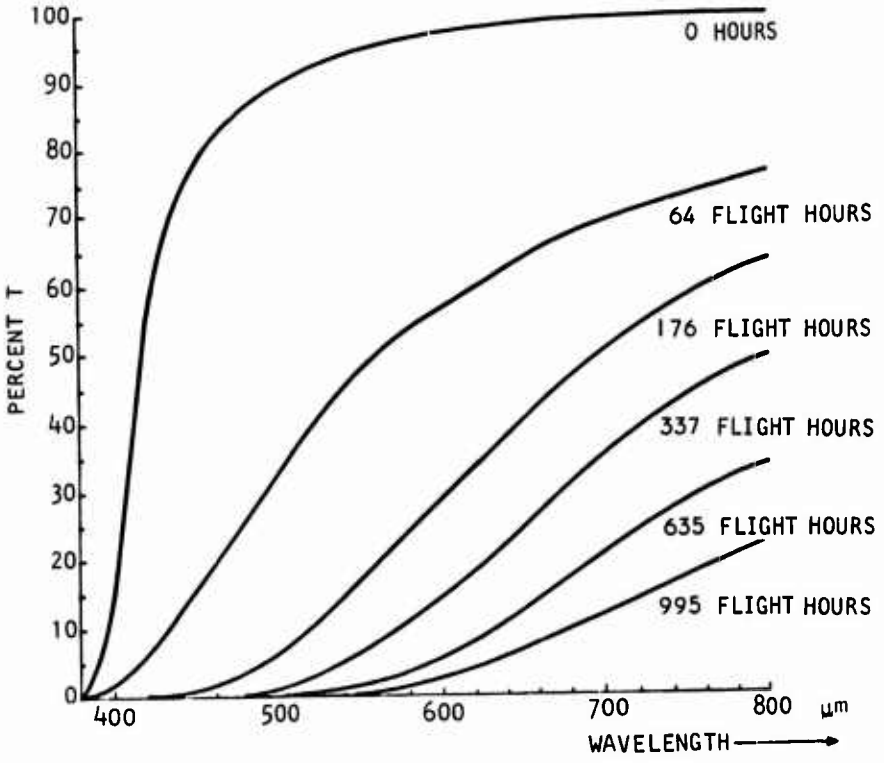



Figure 17. Change of Lubricant Light Transmittance as a Function of Flight Hours. (Figure 2 of Reference 16)

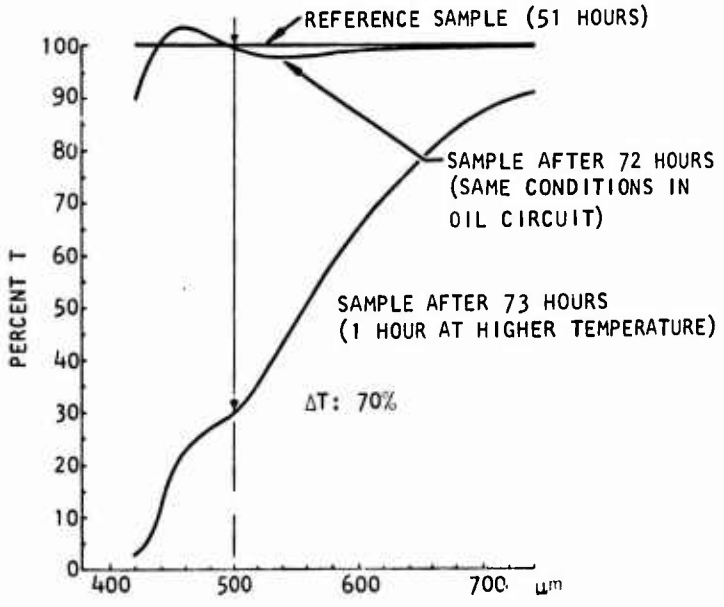



Figure 18. Change of Lubricant Light Transmittance as a Result of Operation at Elevated Temperature. (Figure 6 of Reference 16)

capable of detecting lubricant degradations resulting from approximately 10 to 15 minutes of excessive thermal exposure.

Density

Use of digital lubricant density measurement methods allowing recording to six decimals improved reference scatter range from ±0.0002 g/ml to ±0.000025 g/ml. Typical density changes in approximately 700 hr of in-flight use of lubricant were from 0.921550 g/ml to 0.921950 g/ml, or +0.000400 g/ml.

Elementary lubricant analysis for determination of additive status in the form of the X-ray fluorescense method provides detection limits for chlorine, sulphur, and phosphor better than 0.01 percent by weight, with reproducibility not exceeding 0.005 percent by weight. Infrared spectroscopy as used for the determination of nitrogenous additives allows detection changes in nitrogen content of approximately ±0.01 percent by weight. Infrared spectroscopy is also the most commonly used method to detect water content in lubricant, a frequent cause of component degradation in helicopter transmissions.

Primary Candidates

Of the five lubricant characteristics discussed above, the primary candidates as indicators of lubricant degradation, and thus of the existence of abnormal conditions inside a transmission or gearbox useful for life prognosis, are changes in light transmittance and acid number. It could be recommended that analysis of lubricant physical/chemical conditions (as characterized by the above five primary categories) be performed periodically during a prognostic transmission life expenditure test program to determine quantitatively the lubricant characteristics pattern changes concomitant with the life expenditure process.

## NONDESTRUCTIVE TEST METHODS

The nondestructive test (NDT) techniques currently in use or under development were critically evaluated during the program with respect to on-line, real-time operational prognostic application. A technique is considered to have on-line prognostic application potential if the use of the technique does not require the disassembly of the mechanical system under inspection. A technique is said to have real-time prognostic application potential if the prognostic results can be obtained shortly after the application of the technique. In order to be operational, a technique must yield electrical output and its data acquisition procedures must be machine implementable or automatable, in addition to the on-line, real-time requirements.

Generally speaking, the NDT techniques currently in practice or under development are more or less sensing techniques only. Either directly or indirectly they provide an indication of the number and/or size of the surface, subsurface, or interior flaws through the specific process of the individual techniques.

The evaluation is discussed in detail below. Characteristics of the techniques that are important to the on-line, real-time operational prognostic application are summarized in Table IV. Among the 18 techniques evaluated, the ultrasonic pulse-echo technique appears to be the NDT technique with the greatest potential for prognostic application. This technique is reasonably developed, can be applied to both metals and nonmetals, and can be incorporated into a field type prognostic device. The flaw indication obtained using this technique is reasonably good; its great penetration capability provides the potential for detection of surface, subsurface, and deep interior flaws without system or component disassembly. It is conceivable that an operational pulse-echo type system that provides a 3-D flaw condition information display for a medium-size mechanical system such as a helicopter gearbox can be achieved with electronic scanning and computer reconstruction of the reflected image. Since the system can be designed for quick external attachments of the transduction array to the gearbox, permanent installation is not required, and the device can be shared among a large number of helicopters during the ground inspection periods.

The stress wave analysis technique appears to be the second-best candidate if the noise contamination problem can be prevented. Because of their simplified transmitting and receiving requirements, the eddy current and the Barkhausen-effect techniques can be used for spot-monitoring some of the important and critical components inside the mechanical system. They are not suitable for monitoring assembled systems, however, because their short penetration depths are insufficient to sense all internal component flaws.

Radioactive Category

Radioactive methods involve the use of radiation from either an X-ray or another radioactive isotope source for the inspection of surface and subsurface flaws of a solid mechanical component. Two types of radioactive radiation are being used: (1) the X-ray and (2) the gamma-ray. The X-ray radiation is generated by a heated-cathode-type X-ray tube. The gamma-ray radiation can be obtained from a number of radioactive isotopes and fission products such as radium, cobalt-60, iridium-192, thulium-170, and cesium-137. Because of the short wavelength radiations of this type, the interaction between these radioactive rays and the solid material structures takes place at the atomic and subatomic levels. Therefore, the absorption pattern of the solid structure to the impinging radioactive emissions is the only practical means of detecting macroscopic flaws by this class of NDT techniques.

Although this class of techniques is quite effective in detecting and assessing the flaw, the requirement of bulky protective walls against radiation damage and the inconvenience due to film development make these radioactive techniques poor candidates for on-line, real-time operational types of fault prognosis applications.

According to the type of information presentation, this class of NDT

## TABLE IV. CHARACTERISTICS OF THE NDT TECHNIQUES

| Category | NDT Techniques Evaluated | Technique Status | Applicable Material | Applicable Test Environment | Reliability of Flaw Indication | Detectable Flaws Type | Automation Potential | Monitoring Scope | Special Requirements |
|---|---|---|---|---|---|---|---|---|---|
| Radioactive | Radiography | Reasonably Developed | Metal | Laboratory | Good | Surface, subsurface, and deep interior | Partially machine implementation | Systems and components in assembled form | Needs bulky walls for radiation protection |
| | X-Ray fluoroscopy | | | | | | | | |
| | X-Ray laminograph | Developing | | | | Surface and subsurface | Machine implementable | | |
| Ultrasonic | Ultrasonic pulse-echo | Reasonably developed | Metal and nonmetal | Field and laboratory | Good | Surface, subsurface and deep interior | Machine implementable | Systems and components in assembled form | -- |
| | Ultrasonic resonance | | | | | | | | Regularly shaped test object only |
| | Acoustic holography | Developing | | Laboratory | Unknown | Surface and subsurface | | | Requires bulky reconstruction device |
| | Stress wave analysis | | | Field and laboratory | | | | | -- |
| Magnetic | Magnetic field perturbation | Reasonably developed | Ferromagnetic material | Laboratory | Good | Surface and subsurface | Machine implementable | Components in disassembled form | Requires bulky magnets |
| | Magnetic particle | Developed | | | | | Manual | | |
| | Barkhausen-effect | Developing | | | Unknown | | Machine implementable | | -- |
| | Magnetoabsorption | | | | | | | | |
| Electrical | Injected electric current technique | Reasonably developed | Metal | Laboratory | Good | Surface and subsurface | Machine implementable | Components in disassembled form | -- |
| | Electrified particle | Developed | Nonmetal | | | | Manual | | |
| Electromagnetic | Eddy current | Reasonably developed | Metal | Field and laboratory | Good | Surface and subsurface | Machine implementable | Components in disassembled form | -- |
| | Microwave | Developing | Nonmetal | Laboratory | | Surface, subsurface and deep interior | | | |
| Miscellaneous | Chemical pickling | Developed | Metal and nonmetal | Laboratory | Good | Surface | Manual | Components in disassembled form | -- |
| | Penetrant | | | | | | | | |
| | Laser holography | Developing | | | | | Machine implementable | | |

methods can be further subdivided into three typical groups: (1) radiographic technique, (2) fluoroscopic technique, and (3) laminographic technique. These are discussed below.

Radiographic Technique

The radiographic technique (Reference 17) is essentially a negative shadowgraph produced by the passage of the radioactive rays through the object under examination onto a recording film. The part that allows more rays to pass through exhibits a darker shadow than the part that transmits fewer rays. Because the subsurface or surface type flaws or cracks provide an easier transmission path to the radioactive rays, a dark image of the flaws or cracks can be generated on the film to allow detection. When the radiation source is an X-ray tube, the resultant shadowgraph is usually called the X-ray radiograph. When the radiation source is a radioactive isotope that emits gamma-rays, the resultant shadowgraph is called a gamma-graph.

The penetration depth of these radiographs depends upon the intensity of the radiation source and the geometrical arrangement of the object with respect to the source. The sharpness of the radiograph picture depends upon the size of the emission source and the distance between the source and the photographic film. The contrast of the picture depends upon the source intensity and the exposure time. To obtain a good picture of the flaw, a proper combination of these factors must be employed.

If the depth at which the flaws occur is to be determined, the stereoradiography and the parallax techniques can be used. The stereoradiography technique makes two radiographs of the object with the source moved a distance equal to the interpupilary distance between the eyes, and then views the two radiographs through a stereoptical viewer to get the 3-D transparent picture of the object under examination. The depth of the flaw can be estimated by comparing the dimensions with some known calibration scales. The parallax technique determines the depth of flaws by double exposure. This technique makes two exposures of the object on the same film at two different positions so that the images of the flaws on the film will exhibit a shift from one location to another on the film. Lead markers on the source and film sides of the object are used to allow the quantitative correlation of the shifting movement to some known dimensions. The depth of the flaws can then be calculated based on the estimates of the flaw and marker image shifts.

Normally, the X-ray radiographic technique is used to inspect small or single items, and the gamma-ray radiographic technique is used to inspect large items or a number of items simultaneously by placing the radioactive source at the center of the object to be inspected. These techniques leave a permanent record of the object after the inspection and involve the use of a relatively bulky source and sizable protection walls, or remote operation of the inspection process. They also require the availability of a dark room for loading, unloading, developing,

washing, and fixing during the film processing. The radiographic techniques have been used quite successfully in a test laboratory environment. Because of the need for protection against the radiation damage, however, the operational application of these techniques does not appear to be very promising except for very small objects where the required source intensity is low.

Fluoroscopic Technique

This method (Reference 17) is a variation of the radiographic method discussed above; instead of using a photographic film, a visually observable fluorescent screen is used. This allows the accurate radioactive internal inspection of solid objects at a lower cost, higher volume, and faster speed than the radiographic method. If a permanent record is required, a camera can be used to photograph the image being inspected on the fluorescent screen. A TV camera can be used to allow remote viewing of the image, which somewhat reduces the probability of radiation damage to the viewer. To ensure the safety of other personnel, however, sufficient radiation protection is still necessary even with the remote viewing capability.

A typical radioactive fluoroscope is usually made of two major subunits, i.e., a cabinet and a viewing arrangement. The cabinet is lined with lead material for protection and contains the X-ray generating equipment, a port entrance and exit, the object handling mechanism, a fluorescent imaging screen, and a thick viewing window made of lead silicate glass plates. The viewing arrangement may take the form of a dark room, a booth, a hood, a curtain, or a TV monitor. The object to be inspected is placed on a turntable of the handling mechanism between the radiation source and the viewing screen. Through use of this turntable arrangement, the object can be inspected continously from different angles. The radiation source used in the fluoroscope is usually the X-ray.

Contrary to the radiographic technique, the fluorescent material used on the screen generates a bright image for the part of the object that allows more radiation to pass through instead of a dark one, as in the case of the radiographic technique. The sensitivity of the fluoroscopic technique is generally inferior to that of the radiographic technique because it depends upon the secondary emission of visible light from the fluorescent chemicals. Since the inspector can control the viewing angle, fluoroscopy is particularly useful in more easily locating the linear type of discontinuities, such as a tear or crack. Shrinkage may be seen readily if it is of the cavity or coarse sponge variety; however, microshrinkage is difficult to detect by fluoroscopy. The preferred geometry type inclusion of foreign material can be detected more easily by fluoroscopy than by radiography. Because of the nonuniformity of the shape of the porositic voids, the sharpness of the image of the voids is consequently reduced. The detection of these voids is equally difficult using either fluoroscopy or radiography. The fluoroscopic technique is usually used to perform gross inspections for small objects, while the radiographic technique is

used to provide refined inspections or inspection for large components. The main advantage of fluoroscopy over radiography is the flexibility and the economy of the former technique.

Laminographic Technique

This is another variation of the basic radiographic method of NDT inspection (References 18 and 19). By a special arrangement, this method allows examination of a thin section of a thick object without physically sectioning the test object. X-ray is usually used as the radiation source for this method, possibly because of the flexibility in control. The basic operating principle is averaging (or smearing) an unwanted image over a large area while the image of interest is kept sharp throughout the exposure. This objective is accomplished by synchronously rotating the object under inspection and the recording film during the exposure. Such laminographic equipment includes a point X-ray source, a rotating object table, and a rotating film table. The axes of the two rotating tables are in parallel. The object table is placed between the point radiation source and the recording film table to allow the image of the object to be cast upon the film. Because of both object and film motion, however, only a very thin section (whose image always coincides constructively on the film) of the object under inspection will be enhanced and yield a sharp image on the film. The image of other sections of the object will be averaged out to yield a monotonous background. The particular section of the object to be radiographed can be selected by adjusting the inclination angle of the radioactive ray with respect to the turning tables, the distance between the tables, and the distance between the axes of the rotating tables. This X-ray laminography technique can separate layers that are as close as 0.004 in. and can detect flaws as small as 0.0007 in. To speed up the inspection, an X-ray fluorescent screen can be used instead of the film. The laminograph then can be viewed directly on the screen or indirectly through a TV monitor camera. Because precise synchronization of the two rotating tables is required, the X-ray laminography currently applies to small objects only, such as the NDT inspection of internal flaws of the density packaged multilayer printed circuit board.

Ultrasonic Category

The ultrasonic methods currently either being used or in the process of development for NDT applications can be grouped in the classes of active ultrasonic and passive ultrasonic. For active ultrasonic, the methods employed consist of the well-known and well-developed pulse-echo technique, the continuous-wave resonance measurement technique, and the developing acoustic holography technique. The only method of the passive ultrasonic class is the stress wave analysis technique (SWAT), which is a variation of the acoustic emission technique discussed earlier but using a different data processing approach. The pulse-echo technique, the continuous-wave resonance measurement technique, the acoustic holography technique, and

the stress wave analysis technique are discussed below.

Ultrasonic Pulse-Echo Technique

Probably the oldest ultrasonic NDT technique in use (References 17 and 20), the ultrasonic pulse-echo technique is based on the principle that a reflecting wave is generated every time an incident wave crosses a boundary of different acoustic velocities. Since mechanical flaws, especially the cracks, normally create such a situation, their existence can be detected by analyzing the reflected echo waves from the flaws.

The test setup generally involves an ultrasonic signal source, a transducer that serves as both a transmitter and a receiver, the specimen under test, and a data presentation device. The ultrasonic waves are transmitted in pulses to the test specimen through the transducer. The differential time between the moment that the incident wave is applied to the surface of the specimen and the moment the reflected wave arrives at the transducer (receiver) can be used to locate the flaw, and the intensity of the reflection can be used to assess the size of the flaw. To detect surface type or shallow subsurface type flaws, a wedge coupling device can be used to adjust the incident angle of the ultrasonic beam so that the transmission of the ultrasonic waves in the test specimen can be limited to a thin layer below the surface. To cover all possible flaw locations of interest in a given specimen, scanning of the specimen will be required. This is normally accomplished either manually by test personnel or semi-automatically with some kind of scanning mechanism. This bulky mechanism usually limits the application of the ultrasonic pulse-echo technique to off-field environments. The development of an electronically steerable scanning scheme (Reference 20), however, points to the possibility of shifting the ultrasonic beam without the use of a moving mechanical device even though improvements in transducer array design may still be required. By proper control of the phase angle at the various transducers in an array and by proper sequencing of these phase relationships, it appears possible to steer the ultrasonic beam over a larger angle in a manner similar to the scanning employed in an electronically controlled radar antenna array.

Three types of data presentation schemes are used with the pulse-echo technique: A-scan, B-scan, and C-scan. A-scan displays the ultrasonic pulse-echo amplitudes vertically on a cathode-ray oscilloscope. The horizontal time difference between the transmitted pulse and the reflected echo can be calibrated to give the flaw depth information. The vertical excursion of the echo pulse represents the flaw size. This data presentation provides both the flaw size and location information on a single display. B-scan provides a display equivalent to a cross-sectional view of the test specimen described by the scanning path of the transducer. This presentation is usually made on a photographic film in conjunction with a CRT serving as the light source and a film-moving mechanism that moves in synchronism with the scanning

transducer. The size of the flaw is represented by the size of the blip on the film. The depth of the flaw can be estimated from the distance between the blip and other reference lines that represent either the top or bottom line of the cross section. C-scan produces an image on a memory-type CRT tube or a facsimile recorder corresponding to a plan or a top view of the test specimen. This data presentation usually does not provide the depth information. Although it has not been implemented, it may be possible to provide a 3-D see-through machine drawing of the test specimen together with the flaws if the scanning information and echo are properly handled by a digital computer device. In view of the rapid cost reduction of mini-computers, such a 3-D display also may be economically implementable in the near future.

The data presentation schemes used in the ultrasonic pulse-echo technique are quite effective. They can be modified easily to facilitate the machine decision-making. The sensitivity of the ultrasonic pulse-echo technique is usually adequate. Cracks as small as 0.00002 in. thick can be detected easily at the center of sections 30 to 40 in. thick. The penetration capability of ultrasonic waves in metal is extremely large. Commercial equipment available as off-the-shelf products can penetrate a metal structure as much as 100 ft in length. With the use of surface wave transmission, flaws located at 0.1 in. below the surface can be detected. This pulse-echo technique has been used to detect laps, seams, rolling cracks, laminations, inclusions, and other defects in steel plates 1/4 in. to about 12 in. thick. Discontinuities on the order of 0.5 percent of the plate thickness are readily detectable, as are laminations down to less than 0.00002 in. thick. It also has been used to locate porosity, cupping, pipe, internal ruptures, and nonmetallic inclusions in bar stock and ingots of various sizes up to 48 in. in diameter in a variety of materials--aluminum, zirconium, plain carbon and stainless steels, high-temperature alloys, and atomic fuels, to name a few. Ultrasonics have been used to inspect cylindrical forgings such as turbine spindles and generator rotors. Varying in size from 6 in. to as large as 60 in. in diameter, some of these forgings weigh as much as 100 tons. The ultrasonic pulse-echo technique also has been used to locate cracks, blow holes, insufficient penetration, lack of fusion, and other discontinuities in welds of all types. With special transmission incident angle control, discontinuities equal to 1 percent of the weld thickness in size can be detected. Compared to the radiographic and the fluoroscopic techniques, the ultrasonic pulse-echo technique provides a much clearer image of a flaw, possibly because of less influence of material uniformity on ultrasonic wave transmission than on radioactive ray transmission.

Ultrasonic Resonance Measurement Technique

This technique (Reference 17) requires the use of a continuous-wave-type ultrasonic emission to be forced upon a test object. The frequency of this continuous wave will be varied until at a given value

the distance between the transducer point and the flaw is half wavelength (or integer multiple of half wavelength). Under this condition, the reflected wave from the flaw will be in phase with the incident wave at the transduction point. Therefore, a resonance will be set up at the transduction point. The occurrence of the resonance easily can be detected by the use of a CRO, a meter, or a head phone.

This technique provides information about only the location of the flaw. The intensity of the resonance depends on many factors other than the size of the flaw; consequently, this information cannot be used to indicate the flaw size. The size of the flaw has to be relatively larger than the ultrasonic beam size. To allow easier detection of the flaw, the opposite surface of the test object has to be reasonably parallel to the surface touched by the transducer. Because of these limitations, it has been used mainly as a thickness gage for thin devices or for certain inaccessible places. It also has been used to detect relatively large laminar defects or to locate unbonded zones in sandwich materials.

Acoustic Holography Technique

In general, holography is a new optical technique for the recording of both amplitude and phase information of an interference pattern created by a reference wave and an object wave emanating from the object under examination; the conventional photographic technique records the amplitude (or intensity) information of the object wave only (References 21 and 22). When this holographic recording is illuminated by a wave similar to the reference wave, an inverse process takes place so that the original object can be reconstructed with all the spatial details included (i.e., three-dimensional and all angles of views). The holographic technique can reproduce every spatial detail of the object because it can maintain consistent phase relationships among the various wavelets generated from the object. Consequently, a coherent light source for recording and reproduction is required. Optical holography cannot be performed successfully until a coherent source, such as the laser beam, is developed.

Acoustic holography is similar to optical holography except a coherent ultrasonic acoustic wave source is used instead of a coherent light source. Also, different techniques are used to record the interference pattern and to reproduce the image from the recording. Since ultrasonic wave radiation is normally coherent, or can be made coherent, a phase-consistent interference pattern of the reference wave and the wave passing through the object can be formed. The commercially available ultrasonic transducers usually can meet the power and frequency requirements of the acoustic holography technique with little or no problem.

The most critical aspect of acoustic holography is the method of recording the acoustic pressure wave interference pattern and of reconstructing the image from the recording. Although it is possible

to render film sensitive to pressure waves instead of light waves, a great deal of research effort is necessary before such an approach can be effected. At present, there are three techniques used for recording. The first technique is to use an electrical pressure sensor to scan the interference pressure patterns and then modulate a light source with this electrical signal from the pressure sensor for the exposure of a photographic film. The light source has to move in synchronism with the pressure sensor. Therefore, recording processing for such a method is very tedious and time consuming. Another way of recording the acoustic interference pattern is to produce the interference pattern on the surface of a fluid medium such as water, and then scan this pattern with the use of an optical laser beam for the production of a photographic recording. This recording method, however, requires the immersion of the test object and the ultrasonic transducers into the fluid medium. Efforts to find a dry medium on which the acoustic interference can be formed have been made, and no successful method ever has been reported. The third method for recording the acoustic interference pattern involves the use of a digital computer with a massive storage capability. To do this, the hologram must be divided into sufficiently small areas, and the phase and amplitude information at those areas must be stored in the computer memory to allow retrieval for reconstruction of the image. The reference wave needed for the reconstruction can be added electrically to the stored phase and amplitude data during the information retrieval operation. The advantage of using a digital computer over the previous two methods is the flexibility that the digital computer method offers. For example, weighting of data from different areas of the object may allow enhancement over specific parts. With the digital computer, it is not necessary to operate the data or to set an acceptance level until the test is completed. Also, the previous data can be rechecked, and a different analysis on the previous data can be performed. The presence of a flaw can be indicated by variations in the quantities.

The type of flaw can be assessed by the way one quantity varies in relation to the other. In addition, the other two recording methods require the use of an optical laser source to illuminate the resultant hologram for the reconstruction of the image of the object under examination. The great penetration capability of ultrasonic waves into a metal object gives the acoustic holography technique a see-through capability to examine the deeply buried type of flaws that the radioactive methods and the optical holographic method are unable to detect. The information content regarding flaws is probably greater in the acoustic holography than in the pulse-echo type ultrasonic methods.

Stress Wave Analysis Technique (SWAT)

SWAT is a method of locating fractures inside a structure by analyzing the ultrasonic stress wave emanating from the fracturing process (Reference 18). Unlike the other techniques, SWAT is a passive ultrasonic method that operates on signals created from a growing crack instead of signals sent from an active transmitter. Using strategically

located sensors, SWAT senses the shock waves associated with a growing crack. The location of the crack can then be determined based on the time lag for each sensor receiving the shock wave and the triangulation outcome among the sensors. The crack size can be indicated by the amplitude of the shock wave. This information, however, has not been used extensively in current applications.

Although there was no official identification of SWAT with the acoustic emission technique for prognostic analysis, the signals on which both techniques work are at least of the same origin. The sensor used for SWAT can be any dynamic sensor such as a dynamic microphone, a high-sensitivity dynamic pressure sensor, a strain gage, or a high-frequency vibration accelerometer. The selection of the sensor for a specific application greatly depends upon the environmental requirements of the test facility. This technique has been used by Aerojet-General Corporation to detect fractures on pressure vessels and rocket motor cases. Experience has indicated that a crack opening $10^{-6}$ sq in. in size can release enough energy to be measured, and cracks $10^{-5}$ sq in. in size can be detected easily under laboratory test conditions.

Presently, the main problem with SWAT lies in the interpretation of the content of the shock waves sensed. It is possible there is considerably more information carried by the stress wave than is currently being used for NDT analysis.

## Magnetic Category

Magnetic methods use certain magnetic characteristic changes resulting from the existence of mechanical flaws in the structure for the detection and assessment of the criticality of the flaws. This class of techniques includes the better-known ones, such as the magnetic field perturbation (MFP) technique and the magnetic particle technique, as well as some recent developments such as the Barkhausen-effect technique and the magnetoabsorption technique. The fundamental limitation of this class of techniques is the restriction to ferromagnetic materials only.

### MFP Technique

The fundamental principle of the MFP technique (Reference 23) is based on the fact that the magnetic field distribution inside a ferromagnetic material will be changed by the presence of a flaw. If the flaw is large enough in size and close enough to the surface, the perturbation on the magnetic field will cause more flux lines to diverge from the regular path, which is more or less in parallel with the surface. Therefore, by sensing the existence of those diverging flux lines, the flaw can be located and flaw size assessed.

This technique requires the use of a magnetization device to set up a magnetic field in the material, the use of a sensor to pinpoint the location where an abrupt field perturbation exists, and the use of a demagnetization device to neutralize the remaining magnetism after the NDT examination in order to avoid nuisance in operation.

Normally, a dc magnetic field is employed by the MFP technique. This field can be established by a permanent magnet or a dc-excited electromagnet. The sensors used can be a magnetic compass, an inductive search coil, or a Hall-effect probe. Although the magnetic compass provides a visual indication only, the other two types of sensors provide electrical signals that can be either machine-processed or visually displayed through the use of a CRT scope. The demagnetization device usually requires a design to reverse the field direction successively at reduced intensity.

The use of an ac field could simplify the magnetization and demagnetization requirement; however, the penetration depth of the field into the material is limited because of the skin effect. Consequently, the ac field approach can be used only to detect the surface and shallow subsurface types of flaws. The dc field can penetrate deeply into the material; however, the perturbation of the magnetic field due to a flaw becomes more smoothed out by the adjacent material as the depth of the flaw increases. The flaw detection capability of this technique varies, depending upon the intensity of the magnetization, the geometry of the flaw, the type of flaw, and the sensitivity of the sensor employed. It has been reported that a 0.001-in.-diameter inclusion located 0.001 in. beneath the raceway surface of a ball bearing can be detected successfully with the MFP technique (Reference 23). The drawbacks of this technique are:

1. Applicable to ferromagnetic metals only
2. Requires bulky magnetization and demagnetization devices
3. Requires scanning of the sensor to locate the fault
4. Limited penetration depth even with dc magnetic field

Magnetic Particle Technique

This technique (Reference 17) is a variation of the MFP technique discussed previously. Similar to the MFP technique, magnetization and demagnetization of the test object are required. Field disturbance caused by a flaw is sensed by observing the adhesion pattern of the magnetic particles that are dusted on the test object.

Although simple to use, the sensitivity of this technique is inferior to the MFP technique. The penetration depth of this technique varies depending upon many factors, such as the intensity of the magnetization, the type of magnetization (dc or ac), the size of the section under test, the size of the flaw, the location of the flaw in relation to the section, and the geometry of the section under test. In some instances, a penetration depth of 5/8 in. has been reported. A more reliable penetration depth is probably about 3/16 to 1/4 in.

This technique detects surface type flaws and occasionally, under certain conditions, the subsurface flaws. Common defects such as

shrink cavities, blow holes, nonmetallic inclusions, segregates, laps, bursts, cracks, seams, heat treatment defects, grinding defects, etching defects, machining defects, and sudden-death-type fractures are detected. For instance, it has been reported that a sand pocket 8 in. long by 3/4 in. deep by 1/2 in. wide and lying 1/2 in. below the surface has been detected using this technique. Because bulky magnetization and demagnetization devices are required for this technique, it is limited to use in the NDT type laboratory environment. Also, because the test outcome does not render an electrical output, use of this technique is limited to machine processing of the test results.

Barkhausen-Effect Technique

Essentially, the Barkhausen-effect technique allows the measurement of the applied and the residual stress levels of a section of the test object (Reference 23). Since the stress concentration is both a cause and an effect of a fracture or crack, the ability to assess the stress level on or inside an object allows inference of the mechanical integrity of that object.

The Barkhausen effect is that phenomenon where the magnetic domains inside the ferromagnetic material undergo abrupt jumps instead of continuous variations when the magnetization force is increased or decreased continuously. When these abrupt jumps are sensed by an inductive search coil, a high-frequency, random noise signal (Barkhausen noise) is created. The mean amplitude of this noise signal is a function of the magnetization. It increases with magnetization to a certain extent, and then decays as the test specimen nears magnetic saturation. The plot of this amplitude versus the magnetization indicates that there is an empirical correlation between certain characteristics of the plot with the stress state of the section under measurement.

The two most sensitive characteristics of the curve that reflect the stress state of the specimen are the peak amplitude of the curve and the area under the curve. It has been observed in a laboratory environment on prepared test specimens that both the peak amplitude of the curve and the area under the curve change monotonously with the stress level of the specimen, and the area under the curve follows a more linear relationship. The repeatability of the test results is reasonably good, approximately 3 percent in the elastic range.

Since both the peak amplitude of the curve and the area under the curve can be obtained from the sensed Barkhausen noise signal by electronic means, this technique appears to have the potential for operational, on-line, and real-time measurement of the stress state information (including the presently applied and past residual stresses) that is needed for making the prognostic predictions. Before an operational device can be realized for this purpose, however, a great deal of research effort is required, including (1) refining probe design to improve penetration and resolution, (2) studying the practical schemes for scanning the test object, (3) analyzing the

effects of the geometry of the test section on the accuracy of measurement, and (4) studying the most suitable form of data processing that will economize the cost of the device.

### Magnetoabsorption Technique

The magnetoabsorption technique is based on the principle that the magnetic energy absorbed in a ferromagnetic material in the form of magnetic hysteresis depends upon the stress status that the material has undergone. To measure the change in magnetic energy absorbed, a low-frequency magnetic field is applied to the test specimen so that the hysteresis loop can be traced out. The energy absorbed at each operating point on the hysteresis loop is measured in terms of the reflected electrical parameter changes ($\Delta R$ and $\Delta L$) of an RF induction coil that imposes an additional high-frequency magnetic field to the slow-varying magnetic bias that traces out the hysteresis loop. If the excitation oscillator is properly designed, these electrical parameter changes can create a modulation pattern on the RF excitation voltage and consequently allow the detection of changes of magnetic energy absorption by electrical means.

Unlike the Barkhausen-effect technique, however, the magnetoabsorption characteristics relate to the stress status of the test specimen in a much more complicated manner. In general, the shape of the entire magnetoabsorption curve traced out by the low-frequency bias field is indicative of the stress level. This makes the interpretation of the test results less quantitative, thus making machine interpretation more remote. This technique has been explored only in laboratory environments on specially prepared specimens. Enough data do not exist to evaluate its performance. A unique feature of this technique is its ability to determine the direction and the type (compression or tension) of the stress based on the manner in which the magnetoabsorption changes and the a priori knowledge of the magnetostriction constant of the test specimen.

## Electrical Category

The electrical category technique uses the disturbances in electrical current flow or field distribution caused by a flaw to detect and assess flaw criticality. Two typical categories of techniques belong to this class: (1) injected electric current perturbation technique, and (2) electrified particle technique. These are discussed below.

### Injected Electric Current Perturbation Technique (IECP)

The injected electric current perturbation technique (References 17 and 22) operates on a principle similar to that of the MFP technique except that the working medium in this case is the electric field instead of the magnetic field. A flaw of sufficient size will create a distortion on the flow path of the electric current and consequently cause a change in the electric potential distribution on the test

object. To accomplish this, electric current from either a dc or an ac source is injected into the test object through a pair of contact probes or brushes. The perturbation due to the flaw can be detected by measuring the potential drop across a pair of additional probes placed at equal distances between the two current-injecting probes. The flaw will signify itself as an increase in potential drop as compared to the potential drop of the sound part. This indication, however, is a highly nonlinear function of the thickness of the test section, the probe separation distance, and the depth of the flaw.

To provide enough sensitivity, a large current from a low-voltage source normally is employed. For a large mechanical object such as a turboalternator rotor, a 10,000-amp current has been injected into the test body. A dc source or low-frequency ac source is used if the flaw is deeply buried inside the test object. A high-frequency ac source is usually used to facilitate the sensing and to reduce the current requirement. The flaw depth, however, is limited by the skin effect of the ac current flow.

In addition to using the potential probe method for sensing the electric flow perturbation caused by a flaw, inductive search coils of Hall-effect probes can be used to detect the disruption of current flow by measuring the disturbance associated with the much weaker magnetic field created by this current flow. However, a high-frequency electric current source will be needed; thus a high-gain amplifier and a state-of-the-art electronic method of enhancing the signal-to-noise ratio will be used. The inductive search coil and the Hall-effect probe are nevertheless much more flexible than the contact potential probes. With a 10,000-Hz and 10-amp current source, a surface fatigue crack approximately 0.0003 in. long can be easily detected by this technique with an inductive search coil type sensor. This technique can be applied to both the ferromagnetic and the nonferromagnetic metals. The equipment needed for excitation could be smaller than that needed for the MFP technique. The main limitation of the technique (similar to the main limitation of the MFP technique), however, is the penetration capability to detect deeply buried flaws.

Electrified Particle Technique

The electrified particle technique allows the detection of flaws of a nonmetallic material. Flaws on a nonmetallic object tend to cause charge concentration around them when induced by an external means. The large amount of charged particles being attracted by the charge concentrations will then reveal the flaws themselves automatically, just as in the magnetic particle case. The electrification is obtained by the forced flow of a special type of powder through the hard rubber nozzle of a spray gun. A metal backing or a conducting liquid penetrant is applied to the other side of the test object to allow the free movement of electrons in accordance with the charge concentration. This technique can detect effectively the fine cracks on certain nonmetallic materials such as the porcelain and enamel coatings, glass, fired ceramics, plastics, and some paint films. In certain instances,

fine cracks about $4 \times 10^{-6}$ in. wide have been detected by this technique.

Although this technique is relatively simple to apply, it does not lend itself easily to machine processing and on-line, real-time prognostic testing. Also, since nonmetallic materials rarely are used in mechanical systems such as engines, transmissions, and gearboxes, this technique does not have too much application in the subject program area.

## Electromagnetic Category

The electromagnetic NDT methods utilize the interaction between the flaw and the sensing electromagnetic wave to detect and assess condition status of the flaw. The typical techniques that belong to this class are the eddy current technique and the microwave technique.

### Eddy Current Technique

Essentially, eddy current testing (Reference 24) consists of the observations of the interaction between the electromagnetic wave and the metal in terms of the reflected amplitude and phase information carried in the electric current of a sensing coil. Alternating current is normally injected into the section of the test object underneath the sensing coil through electromagnetic induction. Part of this induced energy is absorbed by the resistivity of the metal, part of the energy is absorbed by the hysteresis of the metal if it is of a ferromagnetic type, and part of the energy is reflected back to the sensing coil through mutual inductance between the coil and the metal. It is this reflected energy that bears the information describing the characteristics, property, quality, and mechanical condition of the object under inspection. This reflected energy appears as part of the current flowing into and out of the sensing coil. With regard to the test object, this condition information can be extracted by electronic means, and the result compared with a standard obtained from a known sound part. The deviation from this standard allows the detection of certain abnormal or flaw conditions.

There are three factors that influence the magnitude and the pattern of eddy current flow in a test specimen; namely, electrical characteristics and properties of the material, the existence of flaw or discontinuous boundaries, and the geometry of the test object. The most dominant characteristic of a metal is its electrical conductivity. Since this characteristic is dependent upon many phenomena on an atomic scale, it is closely related to the strength, hardness, alloy composition, and the heat treatment condition of the test specimen. Also, the conductivity measurement of the eddy current technique allows detection of the differences in these properties.

The existence of flaws or discontinuous boundaries normally interrupts and distorts the flow of current inside the specimen. In an

effort to detour the flaws, the eddy current will experience a phase angle shift. This phase information then can be analyzed by electronic means to assess the size and depth of the flaws.

The geometry of the test object has an overall effect on the magnitude and phase of the eddy current. Because of this effect, the eddy current technique also has been used to determine the dimension or dimensional change of certain mechanical components such as tubes, bolts, ball bearings, and rods. The sensitivity is quite high. Readings accurate to better than $10^{-6}$ in. have been claimed.

Although the effect of various factors that influence the eddy current flow can be theoretically explained, the actual application of the test results is mostly done empirically. Through a period of learning, the inspector then determines the criteria he can use to detect or assess the various abnormalities or properties of the test object. By proper use of phase and magnitude information, an experienced inspector can even identify the cause of a rejectable indication.

This eddy current technique is most suitable for the detection of surface or shallow subsurface types of flaws. It is much more convenient than the MFP technique or the IECP technique because of the relaxed excitation requirement and its high sensitivity. However, the operational application of the NDT technique will still be limited to the more easily accessible areas such as wing structures, fuselage, and landing gear structures of an aircraft because of the skin effect. This technique also is a more likely candidate for automatic machine inspection than the MFP and IECP techniques.

Microwave Technique

The group of microwave NDT techniques (Reference 19) uses the extremely high-frequency (about 1 gigahertz) microwaves as a sensing means to detect the flaws in a test object. This technique is more applicable to nonmetallic materials to complement the ultrasonic and radioactive techniques in testing large objects. Ultrasonic wave energy absorption by the test object of the same size will be unacceptably high. To obtain the same penetration capability by radioactive techniques, much bulky equipment will be required as compared to the microwave system. Although this technique may not find too much application in the subject program area, it is discussed below because it represents a new development in the field of NDT techniques.

Generally speaking, five basic approaches can be employed to use a microwave beam for NDT testing purposes:

1. Continuous-wave (CW) reflectometry

2. Time-domain reflectometry

3. Frequency-domain reflectometry

4. Wavelength-domain reflectometry

5. FM propagation distortion

The CW reflectometry approach is probably the simplest of the various microwave NDT techniques. The CW microwave is sent into the test object and the amplitude of the reflected signal is examined to determine the existence of the flaw. The most detectable flaw using this approach is separation of lamina within the material. Usually, the amplitude of the reflected microwave signal increases with the increase of the separation.

The time domain reflectometry approach is essentially a radar approach. The microwaves are modulated by discrete pulses, and the time difference between the departure of the transmitted pulse and the arrival of the reflected pulse is used to determine the distance between the flaw and the transmitting antenna. The intensity of the reflected pulse provides information with regard to the nature of the flaw. To achieve good resolution at close range, the discrete pulses will be extremely short, even shorter than can be generated in practice.

The frequency domain reflectometry approach is similar to the time domain approach, except the distance between the flaw and the transmitting antenna is obtained by FM means. The present advantage of this approach is that frequency can be measured with greater ease and precision than differential time.

The wavelength domain reflectometry approach is used mainly for cases where closely spaced partial reflective surfaces are involved. This approach uses a swept-frequency generator and a plotting device to plot the relationship between the reflection coefficient and the frequency. When the frequency is such that the spacing of the interfaces becomes 1/4 wavelength, or integer multiples of the 1/4 wavelength, the cancellation will occur and the reflection will be greatly reduced. By searching for the dip in reflection with a varying frequency, the distance between the flaw and the transmitting antenna can be computed.

The FM propagation approach is based on the fact that distortion in the received signal usually contains the information about the condition of the parts to be examined. To permit easier detection (or extraction) of this information, maximization of the distortion is required. Harmonic distortions are analyzed to determine the flaw condition that may exist in the test object.

The microwave NDT technique is relatively new in this field. Although much research is still required, this technique has been used to inspect fibrous composites, laminar structures, solid rocket propellants, and surface cracks on metallic objects. Surface flaws or scratches as small as 100 μin., and separations of about 0.02 in. in a cylindrical section of a glass-filament rocket motor chamber cast with solid propellant, have been detected by the microwave technique.

Miscellaneous Category

Several techniques that are popular in certain NDT applications but lack either the on-line, real-time operational potential or the fruitful prognostic information are discussed below. These are (1) chemical pickling technique, (2) penetrant technique, and (3) laser holography technique.

Chemical pickling is a technique that will effectively remove the metal oxides formed on the surface of a test object to allow a better visualization of the surface condition of that object. This technique generally requires the use of acidic solutions such as hydrochloric or sulphuric acid, or acidic paste, which is a mixture of acidic solution and some holding material such as asbestos wool. The test object will be fully immersed in these chemicals for a period of time, after which the object will be removed, washed down with a water hose, and cleaned with a steel brush to remove all the loose sand and scale. This technique improves the visibility of surface flaws and consequently allows their detection. The flaws detectable by this technique, however, are the large type (approximately 1 inch).

The penetrant technique is based on the idea of using special agents to improve the contrast between the flaw and its background. The test object will be treated with a searching liquid with high penetrating power. After the flaw has been soaked with the penetrant, the surface of the object will be cleaned. The developer then will be applied to the object to encourage the penetrant to emerge from the flaws, showing a dark trace against a white background, a colored pattern against a contrasting background, or self-luminous against a dark background. Apparently neither the chemical pickling technique nor the penetrant technique can be implemented easily by machine to allow automatic operation; in addition, the use of these techniques requires disassembly of the test objects.

Laser holography is an optical method that uses the laser beam as the coherent source needed in the holography. Because of the high attenuation and reflection of the optical rays, this method is limited to the measurement of external dimension changes or the detection of surface flaws.

## PROGNOSTIC IMPLICATIONS OF TRANSMISSION MECHANICAL DESIGN CHARACTERISTICS

### MODULAR TRANSMISSION DESIGN

In the overall failure spectrum of transmission components, the relatively high percentage of secondary failures induced by debris proliferation has contributed to a trend toward subassembly compartmentalization of transmission mechanical design, with provisions for separate lubrication circuits and particle monitoring sensors for individual subassemblies. The implications of this trend on prognostic concepts and possibilities are far-reaching since in such designs abnormal wear particle generation is traceable to specific transmission subassemblies or areas whose mechanical condition changes may influence transmission overall functional integrity or life expectancy in a different manner. In a system consisting primarily of gears and bearings, abnormal but noncritical gear wear particle generation could overshadow critical bearing debris generation. It will be of interest, therefore, to follow transmission modular design developments to see to what degree they will permit determination of the source of observed wear particle generation processes. Another prognosis-related aspect of modular transmission design is the larger number of dynamic signature transmission paths.

### FUNCTIONAL REDUNDANCY OF COMPONENTS OF LIFE-CRITICAL SUBASSEMBLIES

Examples of this redundancy would be the use of two unidirectional thrust bearings in the input quill assembly of some transmission designs, or the provision of collars on mast shafts to prevent axial shaft dislocations in cases of shaft support bearing failure, as shown in Figure 15 of Reference 25.

For transmission designs without redundant features to increase their functional emergency capability, relatively greater prognostic emphasis must be placed on nonredundant life-critical components than on the same components in transmission designs with redundancy features.

### BASIC LOAD CONCENTRATION

Examples of differences in basic load concentration would be in the number of idler gears in planetary subassemblies or in the number of bearings used in drive pinion assemblies. For instance, in the case of the UH-1 transmission, the upper planetary subassembly with its eight idler gears is less life-critical from a viewpoint of basic load concentration than the lower planetary subassembly containing only four idler gears.

## SPECIFIC LOAD CONCENTRATION

In deciding on the relative degree of prognostic emphasis to be placed on individual transmission areas, component design load levels have to be considered. In the case of gears, bending and compressive stresses (in the case of bearings, roll loads, and Hertzian stresses) can be used as overall measures of specific load concentration.

In view of study objectives, the design concepts related to gear teeth surface durability (as formulated in AGMA 215.01) are of interest:

In the surface durability power formula, which contains a life factor $C_L$,

$$P_{ac} = \frac{n_p F}{126{,}000} \cdot \frac{I\ C_v}{C_s\ C_m\ C_f\ C_o} \cdot \left[ \frac{s_{ac} d}{C_p} \cdot \frac{C_L\ C_H}{C_T\ C_R} \right]^2 \qquad (15)$$

where $P_{ac}$ = allowable power, hp

$n_p$ = pinion speed, rpm

$F$ = net face width, in.

$d$ = pinion operating pitch diameter, in.

and these design characteristics are correlated to the life factor by the following ten factors:

$I$ = geometry factor
This factor relates to pitch line loading and as such would be influenced by operational gear mesh geometry changes due to tooth flank wear or change of alignment conditions.

$C_v$ = dynamic factor
Apart from mesh geometry changes, changes of lubricant properties can influence this factor invoking tribological considerations.

$C_s$ = size factor
This factor reflects the effect of dimensions on the uniformity of material properties. Three subfactors (area of contact pattern, ratio of case depth to tooth size or hardness pattern, and hardenability and heat treatment of materials) could be subject to change during operation or only approximately definable during the design stage.

$C_M$ = load distribution factor
This factor is influenced by internal bearing clearance, shaft parallelism, bearing deflection, thermal expansion, and distortion due to operating temperatures. All of these conditions may deviate in actual operation from their designed-assumed character.

$C_f$ = surface condition factor
This factor will be strongly influenced by gear flank wear conditions.

$C_o$ = overload factor
It can be assumed that, in the case of a specific transmission installation, the operational normal and abnormal steady-state and transient load conditions are sufficiently established to permit a realistic assumption of this factor.

$C_p$ = elastic coefficient
Aside from temperature effects, this coefficient (a function of pinion and gear Poisson's ratio and modulus of elasticity) does not contain any primary elements that would be subject to change in the installed, operational use mode compared to design assumptions made.

$C_H$ = hardness ratio factor
This factor is a function of the pinion-to-gear hardness ratio and as such may be subject to operationally-caused changes.

$C_T$ = temperature factor
Since this factor is assumed at unity for gears operating with oil or gear flank temperatures not exceeding 250°F, operation at higher temperature levels would change this factor.

$C_R$ = factor of safety
This factor, representing an overall design criterion for high reliability or calculated risk, is of no immediate usefulness for diagnosis/prognosis.

The allowable contact stress number, $S_{ac}$, is of interest since it depends, among others, on the number of load cycles. For each material and for the minimum surface hardness and hardening treatment, an allowable contact stress number for 10 million cycles of load application is determined empirically.

The life factor, $C_L$, finally adjusts the allowable loading for the required number of cycles. AGMA 215.01, while providing a curve of life factor vs required life in cycles for pitting, makes only cursory statements regarding this factor.

AGMA Information Sheet No. 217.01 on gear scoring contains prognostically valuable data on this particular gear degradation/failure mechanism. The flash temperature index (in °F) indicative of scoring probability is defined as a function of initial or oil inlet temperature, effective tangential load, effective gear face width, surface finish (after run-in), a scoring geometry factor, pinion rpm and transverse, and operating diametral pitch.

For comparing the scoring tendency of similar gear designs, the scoring index concept used is a function only of effective tangential load,

effective face width, pinion speed and transverse, and operating diametral pitch. Temperature, surface finish, and scoring geometry factors are not included in this concept.

## LUBRICATION MODE

The influence of the transmission and gearbox lubrication mode on prognostic concepts depends upon the extent of particle content and temperature control; for instance, circulating pressure-lubrication systems tend to attenuate the effects of an abnormal component condition, whereas splash lubrication methods as employed for gearboxes do not. In splash-lubricated systems, lubricant condition in terms of particle content, temperature, or physical-chemical characteristics (such as viscosity and acidity) therefore could be expected to be more influenced by an abnormal mechanical component condition than in pressure-lubricated systems. The fact that failure incidence rate and lubricant viscosity, or acidity, were not correlatable in the investigation performed by Bowen (Reference 15) may be at least partially so explained. The accuracy and reproducibility of the type of lubricant analysis used could have been another contributing factor.

## TRIBOLOGICAL CONDITIONS

For gears as well as for bearings of a specific transmission design type, differences in tribological conditions such as elasto-hydrodynamic film thickness, gear conjunction temperature rise, sliding-to-rolling ratios, and surface texture of individual transmission components, must be established in order to place prognostic emphasis on such components whose operational tribological conditions are relatively unfavorable. As recent mode of failure investigations of helicopter transmissions have shown (Reference 15), the present conceptual definition of these conditions appears to describe actual conditions in a manner resulting in incompatibilities with observed component life expenditure rates. Since high bearing failure rates were observed at EHD film thickness three to four times greater than composite surface roughness, a condition for which bearing life in excess of AFBMA guidelines could have been expected, these more traditional life criteria may assess minimum film thickness inadequately by neglecting the effects of temperature rise in sliding contact. As noted by Bowen, another factor contributing to erroneous estimates of elasto-hydrodynamic film thickness appears to be the scarcity of information on the pressure viscosity coefficient entering into the EHD film thickness formulas. With regard to surface finish as well as surface texture, the important role of surface texture in the process of rolling contact failure has been realized only recently. The observation that random contact of surface asperities leads to a more rapid development of fatigue cracks than repeated, localized contact is also significant for life expenditure phenomenology. As stated in the discussion of P. H. Dawson's 1969 paper on rolling contact fatigue crack initiation (Reference 26), use of a lubricant containing certain EP additives could reduce pitting life by as much as one order of magnitude.

With regard to transmission gear life expenditure, the above-referenced failure mode investigations also showed that two gears operating under comparably unfavorable tribological, or FHD film thickness, conditions had a widely different life expenditure rate reflected in surface-initiated pitting. It was concluded therefore that critical differences in lubricant state must have existed. Although these differences were not established, and micro-structural material analysis of the pitting condition (facilitated by use of the scanning electron microscope) afforded deeper insight into the mechanism of the pitting process, the causes are still vague. This situation stresses the need for closely monitoring current tribological research results for background information on transmission life expenditure processes to be followed by a prognostic methodology.

Experience gained during the development of a turboprop engine transmission and discussed by Stewart and Hollingsworth (Reference 27) gives valuable insight into the combined influence of design parameters and operating conditions on the mechanical integrity of transmission components. During development of the PT6A engine reduction gearbox, it was observed that the intensity of a cavitation condition causing erosion of the first-stage planet gear bushings was strongest at a gearbox vibration frequency of 8,000 cps. Changing the meshing frequency of the first planetary stage of 12,000 cps reduced the cavitation intensity and, in consequence, resulted in significantly reduced secondary damage effects. Knowledge of the existence of a cavitation condition as well as of its character and duration within a typical transmission service use period (a now attainable diagnostic load-history task) could represent a major input to prognostic decision making.

## FAILURE MODES AND PATHS

Review of transmission and gearbox component failure modes shows that (1) bearing failures are the prevalent transmission failure modes, (2) secondary failures caused by debris propagation and retention are the prevailing failure modes of bearings, and (3) the type of subsurface-initiated failure treated in traditional bearing life determination practice occurred only in a minority of bearing failure cases, surface-initiated failure being the dominant failure mode.

Since most gear failure modes appear to be of the surface-initiated spalling type, current life-predictive concepts based on subsurface compressive or Hertzian stresses are not valid for the majority of cases.

The evaluation of transmission/gearbox failure studies and statistics provides little information on failure paths. There are indications, however, that operational interactions between dimensionally unstable structural transmission components and rotating components, not adequately realized or allowed for in the design phase, often play a significant role in that area.

Among potential failure paths in a transmission, the interaction between condition changes or degradations of the components of transmission input quill assemblies (such as bearings and spacers), and the input level gear meshing geometry, appears to be of specific interest, even if observed degradations of input quill bearings reportedly rarely progress to a point that changes the static quality of quill shaft support and positioning. Surface degradations of quill shaft support bearings, however, could change the dynamic signature of the input bevel gear mesh significantly by superimposing the abnormal dynamic signature caused by those surface degradations on that of the input gear mesh. If this condition could be differentiated from an abnormal mechanical condition of the mesh itself, which from a transmission life viewpoint may be less critical than an abnormal quill shaft support bearing condition, an important prognostic gain would be obtained. Dynamic analysis experimentation with suitable selected implants of degraded quill shaft support bearings and an as-new bevel gear mesh, as well as with quill shaft support bearings in the intact condition and a degraded bevel gear mesh, would be needed to clarify this point.

Failure modes of helicopter transmissions resulting from abnormal operating conditions (such as interruption of lubricant flow, total loss of lubricant due to battle damage, and lubricant aqueous condensate dilution by water resulting in corrosion of components) have characteristics different from those of failure modes occurring under normal operating conditions. Upon loss of lubricant, bearings must operate under increased preloads that lead to increased heat generation, whereas gears can be expected to remain initially unaffected. As pointed out by Tallian (Reference 28), if temperatures do not stabilize on that increased level, surface wear and galling will occur. Prolonged operation will lead to softening or oxidation of materials and, finally, failure by plastic flow and galling. Prognostic

approaches have to be comprehensive enough to cover these short-term component life expenditure processes caused by abnormal operating conditions as well as the long-term processes taking place under normal operating conditions.

In view of prognostics, the more recent experimental investigations of P. H. Dawson on the origin and progression of the mechanism of fatigue crack initiation are of substantive interest (Reference 26). As Dawson found in his electron microscope studies of fatigue in rolling contact, the first damage indications were detected at less than 5 percent of total life. At a magnification of 400, cracks become visible at 14 percent of the time lapse until they become small pits. At a magnification of 1100, the same cracks could be detected somewhat earlier at 12.5 percent of total life. Dawson concluded that cracks extending to the surface may exist even earlier, and subsurface cracks considerably earlier. An active or corrosive lubricant, of course, would promote such initial cracking substantially.

In the discussion of Dawson's paper by P.S.Y. Chu, the probability of cavitation phenomena representing a contributory factor to fatigue crack propagation is pointed out. Referring to Dawson's experimentation with synchronized as well as freely rolling discs to large-order pitting at a life of 366,000 revolutions, Chu noted that following an initial chemical erosion of the metal surface around the grain boundaries occurring at 214,000 revolutions (or 58.5 percent of macroscopic pitting life), a type of cavitation erosion took place. The same phenomenon was observed in 1959 by A. Bartel and E. Kuss (Reference 29). Even from an initially flat, smooth, solid surface, the sudden release of a high fluid pressure produces these observed types of microeruptions. It was also pointed out that the electronic microscopic studies seem to indicate that plastic flow takes place over the first 124,000 revolutions and that at 244,000 revolutions a deeper and more serious type of pit is formed. In appearance, this pit looks very similar to that caused by an electrostatic discharge across an oil film. Between 250,000 and 258,000 revolutions, a general weakening of the surface around this pit appeared to have occurred. Chu also points out the possibility that initial subsurface and parallel-to-surface cracks may originate even within the first 1000 revolutions.

With lubricant-contained extreme pressure additives, the pitting life may be reduced by as much as one order of magnitude if surface cracking should appear at a life less than 10 percent of the final life-to-failure (based on use of a straight mineral oil).

With regard to failure modes of rolling contact bearings, fatigue failure progression of roller, or line-contact, bearings from initial fatigue spalling to termination of useful life (life margin) will be slower, or life margin longer, than that of ball (or point-contact) bearings due to the higher degree of contact stress concentration in the latter type of bearings. Since some ball bearing applications in helicopter transmissions have had life margins in the order of minutes, an extremely fast response capability of prognostic schemes, or at least a very high early alarm capability, is required. As discussed by Harding (Reference 25), certain

critical transmission component failure modes, in contrast to those of bearings and gears, frequently are not sufficiently realized. These failure modes are gear spline, or disc attachment bolt, fretting, excessive wear of shaft splines and sprag clutches, housing cracking, seal degradations, and mechanical deteriorations of bearing-retaining spacers and liners to the point of annihilating bearing retention.

From a prognostic viewpoint, it is significant that a progressive fretting wear condition has been reported to be generally not detectable by lubricant particle content monitoring methods of the more conventional type. Wear of oscillating linear motion mechanisms (such as shaft splines and clutches) reflects in more accentuated lubricant particle contamination conditions that make these wear conditions detectable. Since abnormal, undetected or undetectable, mechanical condition changes of that type may progress to a critical transmission failure initiating condition, the need for new approaches providing prognostic coverage of these conditions is obvious.

Housing cracks of the progressed type are detectable by lubricant leakage. If dynamic sensors monitoring transmission internal components could be placed in the vicinity of crack-prone housing sections, early crack detection by housing resonance changes might represent a possibility deserving exploration. High wear rate of bearing retaining components is potentially critical. An indicator for a progressive state of this abnormality would be the changes of the character of the dynamic signature of the shaft (a frequency of 1 per revolution) whose guidance is reduced by an excessive wear condition of a bearing retaining component.

## RELIABILITY/STATISTICAL LIFE PREDICTION

Even if statistical and deterministically prognostic life prediction methods have no immediately visible interfaces, the former methods could provide a reference frame for the latter. Also, mathematical modeling techniques of quantitatively formulating and correlating the factors and phenomena determining component life could be useful for prognosis.

## BEARING FATIGUE LIFE MODELING

For a rolling element, the counterpart of the S-N curve for a given material is quite complex. It contains material properties as well as design-configurational properties that are not covered by the stress concentration factor, $K_t$, or other notch effect corrections. For instance, the probability of survival of a ball, $S_b$ is given by

$$\log \frac{1}{S_b} \propto \frac{\tau_o^c N^e a\ell}{Z_o^{h-1}} \qquad (16)$$

where $S_b$ = probability of survival for ball elements

$\tau_o$ = maximum orthogonal shear-stress

$\tau_o$ = C(k) X maximum surface pressure, $p_o$

C(k) = $\tau_o$/maximum surface pressure, $p_o \approx 0.5$ to 0.3

k = a measure of contact area = b/a, the ratio of the minor axis to major axis of the pressure ellipse $= \sqrt{(t^2-1)(2t-1)}$

N = number of $10^6$ cycles

a = major axis of pressure ellipse

$\ell$ = length of rolling path $\approx 2\pi$ x pitch - diameter

$Z_o$ = depth of the maximum orthogonal shear-stress $= b/(t+1)\sqrt{2t-1}$

c,e,h = exponents reflecting material properties of the given bearing

The probability of survival of a roller, $S_r$, is given by

$$\log \frac{1}{S_r} \approx \frac{\tau_o^c N^e \ell_e \ell}{Z_o^{h-1}} \qquad (17)$$

where c,e,h = constants for conventional bearing materials with values for point contact (balls) of

$c = 31/3$

$e = 10/9$

$h = 7/3$

and for line contact (rollers) of

$c = 31/3$

$e = 9/8$

$h = 7/3$

$\ell_e$ = length of roller, and the other quantities are defined in the same manner as those of Equation (16).

The variation of C(k) vs k is depicted in Figure 13.45 of Reference 30.

It should be noted that the maximum surface pressure $\rho_o$ is the maximum Hertz contact stress, but not the tensile stress in a standard S-N curve. It is related to the tensile stress by some material properties (e.g., $\sigma_{t\ max} = \rho_o \left[ (1 - 2\nu)/3 \right]$) when the contact ellipse is a circle and $\nu$ = Poisson's ratio (Reference 31). In view of this, it is clear that either the S-N curves can be used for contact stresses (Reference 32) or the detailed stress states calculated in the contact zone; that is, derive the combined normal and tangential loads and arrive at an equivalent tensile stress,

$$S = f(\rho_o; \nu, \mu)$$

$$= \sqrt{\left(k_1 \frac{1-2\nu}{3}\right)^2 + (k_2 \mu)^2}\ \rho_o = \varkappa \rho_o \qquad (18)$$

where $\mu$ = sliding coefficient of friction

$k_1, k_2$ = weighting factors determined empirically

In order to evaluate C(k) accurately, to convert $\tau_o$ to $\rho_o$, the major and minor axes, a and b, of the pressure ellipse must be obtained from Equations (19), (20), (21), (22), and (23).

$$a = m \sqrt[3]{\frac{3}{4} \frac{P\Delta}{A}} \tag{19}$$

$$b = n \sqrt[3]{\frac{3}{4} \frac{P\Delta}{A}} \tag{20}$$

$$\Delta = \frac{1 - \nu_1^2}{E_1} + \frac{1 - \nu_2^2}{E_2} \tag{21}$$

$$A = \frac{1}{2} \left( \frac{1}{R_1} + \frac{1}{R_1'} + \frac{1}{R_2} + \frac{1}{R_2'} \right) \tag{22}$$

$$B = \frac{1}{2} \left[ \left( \frac{1}{R_1} - \frac{1}{R_1'} \right)^2 + \left( \frac{1}{R_2} - \frac{1}{R_2'} \right)^2 + 2 \left( \frac{1}{R_1} - \frac{1}{R_1'} \right) \left( \frac{1}{R_2} - \frac{1}{R_2'} \right) \cos 2x \right]^{1/2} \tag{23}$$

where P = total radial load on the rolling element

$\nu$ = Poisson's ratio

E = Young's modulus

R = minimum radius of curvature

R' = maximum radius of curvature

x = the angle between the planes containing curvatures $\frac{1}{R_1}$ and $\frac{1}{R_2}$

1,2 = subscripts denoting the two contacting bodies (rolling element and race)

m,n = functions of $\cos^{-1}$ (B/A) tabulated in Table V

TABLE V. VALUES FOR CONSTANTS m AND n

| $\cos^{-1}(B/A)$ | $30^0$ | $35^0$ | $40^0$ | $45^0$ | $50^0$ | $55^0$ | $60^0$ |
|---|---|---|---|---|---|---|---|
| m | 2.731 | 2.397 | 2.136 | 1.926 | 1.754 | 1.611 | 1.486 |
| n | 0.493 | 0.530 | 0.567 | 0.604 | 0.641 | 0.678 | 0.717 |
| $\cos^{-1}(B/A)$ | $65^0$ | $70^0$ | $75^0$ | $80^0$ | $85^0$ | $90^0$ | |
| m | 1.378 | 1.284 | 1.202 | 1.128 | 1.061 | 1.000 | |
| n | 0.759 | 0.802 | 0.846 | 0.893 | 0.944 | 1.000 | |

With all the quantities in Equations (16) and (17) determined, they can be rewritten in the widely adopted formula for S-N curves:

$$N(s)\beta_{b,r} = \nu_{b,r} \tag{24}$$

where the constants $\beta_{b,r}$ and $\nu_{b,r}$ (originally for material properties) are made to be sensitive to the geometric designs of the bearings. This is obvious from a comparison of Equation (24) to Equations (16) and (17). It is clear that $\beta_{b,r}$ and $\nu_{b,r}$ may be expressed as

$$\beta_b = \frac{c}{e} = \frac{31}{3} \cdot \frac{3}{10} = 9.3 \text{ for balls} \tag{25}$$

$$\beta_r = \frac{c}{e} = \frac{31}{3} \cdot \frac{8}{9} = 9.19 \text{ for rollers} \tag{26}$$

$$\nu_b = \left\{ \frac{\left(\log \frac{1}{S_b}\right) Z_o^{h-1} \left[\nu/C(k)\right]^c}{a\ell} \right\}^{\frac{1}{e}} \text{ for balls} \tag{27}$$

$$\nu_r = \left\{ \frac{\left(\log \frac{1}{S_r}\right) Z_o^{h-1} \left[\nu/C(k)\right]^c}{\ell_e \ell} \right\}^{\frac{1}{e}} \text{ for rollers} \tag{28}$$

All quantities at the right-hand side of Equations (25) through (28) are defined earlier. They can be classified into three distinct groups:

1. Material properties: c, e, h and $\nu$
2. Geometric characteristics: $\ell$, $\ell_e$

In order to evaluate C(k) accurately, to convert $\tau_o$ to $\rho_c$, the major and minor axes, a and b, of the pressure ellipse must be obtained from Equations (19), (20), (21), (22), and (23).

$$a = m \sqrt[3]{\frac{3}{4} \frac{P\Delta}{A}} \tag{19}$$

$$b = n \sqrt[3]{\frac{3}{4} \frac{P\Delta}{A}} \tag{20}$$

$$\Delta = \frac{1 - \nu_1^2}{E_1} + \frac{1 - \nu_2^2}{E_2} \tag{21}$$

$$A = \frac{1}{2} \left( \frac{1}{R_1} + \frac{1}{R_1'} + \frac{1}{R_2} + \frac{1}{R_2'} \right) \tag{22}$$

$$B = \frac{1}{2} \left[ \left( \frac{1}{R_1} - \frac{1}{R_1'} \right)^2 + \left( \frac{1}{R_2} - \frac{1}{R_2'} \right)^2 + 2 \left( \frac{1}{R_1} - \frac{1}{R_1'} \right) \left( \frac{1}{R_2} - \frac{1}{R_2'} \right) \cos 2x \right]^{1/2} \tag{23}$$

where P = total radial load on the rolling element

$\nu$ = Poisson's ratio

E = Young's modulus

R = minimum radius of curvature

R' = maximum radius of curvature

x = the angle between the planes containing curvatures $\frac{1}{R_1}$ and $\frac{1}{R_2}$

1,2 = subscripts denoting the two contacting bodies (rolling element and race)

m,n = functions of $\cos^{-1}$ (B/A) tabulated in Table V

3. Derived quantities that are a combination of both material and geometric quantities and the given load: a, b, $Z_o$, k

The quantities $S_b$ and $S_r$ are parameters based on design requirements ($S_b$ or $S_r$ = 1 - accepted failure probability).

With the modified material-geometric S-N curve for bearing established, the Palmgren-Miner theory for cumulative fatigue damage (Reference 33) can be employed to evaluate the expected total damage if the applied stress process is known, or specifically, the n(s) in the following equation:

$$D = \sum_s \frac{n(s)}{N(s)} \tag{29}$$

where n(s) = number of applied stress cycles at stress level s

N(s) = number of allowed stress cycles at stress level s computed from Equation (24)

It is clear that N(s) is deterministic if one is permitted to overlook the fact that the material properties and geometric dimensions are random. Bearing in mind this assumption, the probabilistic part of the damage will be solely dependent on the applied stress process. In brief, the expected damage may be obtained by the expression below:

$$E[D] = \sum_s E\frac{n(s)}{N(s)} \tag{30}$$

$$= \sum_s \frac{E[n(s)]}{N(s)}$$

where $E[\cdot]$ denotes the mathematical expectation of the quantities inside the bracket

By combining Equations (24) and (30), Equation (30) may be changed to the following:

$$E[D] = \kappa^{-1} \sum_s E[n(s)]\, s^{\beta} \tag{31}$$

where subscripts b, r are dropped for simplicity. This expression is general for all stress processes with $E[n(s)]$ either described theoretically by known probability distributions or counted experimentally from actual stress time histories. The latter can be obtained readily from the exceedance curve obtained by peak counting. A discussion of the former (theoretical) approach is presented in the following paragraphs.

Let D(t) and n(s,t) denote two time-parametered random variables representing the damage and number of applied stress cycles respectively. Also,

the stress levels are assumed to extend to $(-\infty, \infty)$ with many discrete bands in between. Then Equation (31) can be expressed by the following:

$$E\left[D(t)\right] = \nu^{-1} \int_{-\infty}^{\infty} s^{\beta}\, ds \int_{0}^{\infty} m\, P_s\,(s,\ t/m)\ P_{M_t}\,(m,t)\, dm \qquad (32)$$

where $m$ = the number of stress peaks

$P_s\,(s,\ t/m)$ = the conditional probability density of the stress peak magnitude, given that the total number of peaks per unit time is $m$

$P_{M_t}\,(m,\ t)$ = the probability density for the total number of peaks per unit time

The conditional probability density $P_s\,(s,\ t/m)$ and the probability density $P_{M_t}\,(m,\ t)$, however, are not easy to estimate for a random stress time history. Engineering approximations must be applied. The second integral may be reduced to the comparatively simple one delineated below:

$$\int_{0}^{\infty} mP_s\,(s,\ t/m)\ P_{M_t}\,(m,t)\, dm = E\left[M_t(t)\right]\ P_s\,(s,t) \qquad (33)$$

if it is assumed that the total number of peaks per unit time is independent of time as well as the stress level s at time t, or

$$P_s\,(s,\ t/m)\ P_{M_t}\,(m,t) \approx P_s\,(s,t)\ P_{m_t}\,(m) \qquad (34)$$

Substituting Equation (33) into Equation (32),

$$E\left[D(t)\right] = \nu^{-1}\ E\left[M_t\,(t)\right] \int_{-\infty}^{\infty} s^{\beta}\ P_s\,(s,t)\, ds \qquad (35)$$

Then, the expected total damage accumulated from time $t = 0$ to $T$ is

$$E\left[D(T)\right] = E\int_0^T D(t)\ dt \tag{36}$$

$$= \int_0^T E\left[D(t)\right]\ dt$$

$$= \nu^{-1}\int_0^T E\left[M_t\ (t)\right]\int_{-\infty}^{\infty} s^{\beta}\ p_s\ (s,t)\ ds\ \ dt$$

The above expression is general and contains no restrictions concerning the random stress process; it can be nonstationary, non-Gaussian, and wide-band. If the random stress is stationary, then

$$E\left[D\ (t)\right] = E\ \ D\int_0^T dt \tag{37}$$

$$= E\left[D\right]\cdot T$$

$$= \nu^{-1}\ E\left[M_t\right]\ T\int_{-\infty}^{\infty} s^{\beta}\ P_s\ (s)\ ds$$

It must be noted that $P_s(s)$ in Equation (37) is only a stationary probability density for peak magnitudes and is not limited to any particular distribution. For example, if the random stress is narrow-band Gaussian, then the peak distribution is Rayleigh and a closed-form solution for $E\left[D\right]$ from Equation (37) is possible. It is given below:

$$E\left[D\right] = \nu^{-1}\ E\left[M_t\right]\int_{-\infty}^{\infty} s^{\beta}\ P_s\ (s)\ ds \tag{38}$$

$$= \nu^{-1}\ E\left[M_t\right]\int_0^{\infty} s^{\beta}\ \frac{s}{\sigma_s^2}\ e^{-\frac{s^2}{2\sigma_s^2}}\ .ds$$

$$= \pi^{-1} E[M_t] (\sqrt{2}\ \sigma_s)^{\beta} \Gamma\left(\frac{\beta+2}{2}\right)$$

where $\Gamma[\cdot]$ is the Gamma function given by

$$\Gamma(x) = 2 \int_0^{\infty} y^{2x-1} e^{-y^2} dy \qquad x > 0 \tag{39}$$

If it is agreed to accept the approximation that for a narrow band the total number of peaks is the same as the number of zero crossings with positive slopes, then Equation 38 can be written as

$$E[D] = \pi^{-1} E[N_o] (\sqrt{2}\ \sigma_s)^{\beta} \Gamma\left(\frac{\beta+2}{2}\right) \tag{40}$$

In passing, it is interesting to note that both quantities $E[N_o]$ and $\sigma_s$ are obtainable by PSD (power spectral density) methods or exceedance curve.

The advanced mathematical model of spalling fatigue failure being developed by Tallian (Reference 28) is based on the assumption that life failure of a rolling contact bearing is preceded by either one or a composite of the following three defect modes: (1) subsurface defects, (2) surface imperfections, and (3) surface fatigue micropits.

For each of these major defect modes, a set of equations quantifies the relevant influence factors. For instance, for the case of the subsurface defects crack initiating mode, the following relationships are used:

1. Crack propagation rate (a function of instantaneous crack size, plastic strain, and ductility)
2. Life at defect site (a function of local plastic strain and ductility)
3. Manson-Coffin life (for low cycle fatigue)
4. Plastic strain (a function of macrostress, hardness, defect size, defect-to-contact area size ratio, and shape/constitution characteristics of defect
5. Ductility (a function of hydrostatic stress and hardness)
6. Deterministic life of a defect of known size and shape
7. Probability of rolling element survival
8. Weibull form of rolling body life distribution

By use of mathematical modeling techniques of this type, it was found possible to correlate the life of a standard bearing operated under typical bearing life test conditions with the life of an in-service bearing manufactured of another type of material and subjected to the particular unfavorable installation conditions of the in-service application. Change factors applicable to the $L_{10}$ life of the standard bearing were as follows:

1. To account for the 20-percent-lower ductility of the in-service bearing material, a change factor of 1/1.8 was predicted.

2. The higher compressive residual stress in the in-service bearing was predicted to account for a change factor of 1.6. (Case hardening of the in-service bearing as compared to the through-hardened condition of the standard bearing was the cause for this higner residual stress.)

3. Large inclusions found in the material of the in-service bearing were predicted to change its $L_{10}$ life by a factor of 1/2.7.

4. The installation conditions of the in-service bearing were found to be such that edge pressures were two times as high as the equivalent uniform pressure. This condition was predicted to account for a change factor applicable to $L_{10}$ life as large as 1/144.

5. Deflection of the support structure of the in-service bearing, causing Hertzian pressure changes in the order of +10 percent over one-half of the bearing outer ring and -10 percent changes over the other half of the ring compared to the Hertzian pressure value for the nondeflected support structure, was predicted to account for a change factor applicable to $L_{10}$ life of 1/144.

The above modeling technique covers life only in the sense of the time period between crack initiation and crack growth to critical size. Prediction of life remaining (i.e., of the time margin between the crack having increased to critical size and the detachment of a spall) is not included in that technique.

It is of interest, in view of component life prognosis, that advanced instrumentation methods such as the shock pulse method, recently developed for bearing condition monitoring (Reference 3), would permit following observationally the progression rate of a spall to critical size, or the end of remaining life. At the point where theoretical life prediction capability ends, instrumentational techniques could be used for monitoring a life expenditure process of the fatigue type in view of failure prognosis.

## PROGNOSTIC METHODOLOGY

Regardless of origin or processing mode, diagnostically obtained information on the mechanical status of a system (a major tributary in a prognostic methodology) can be classified as (1) single-point time history data, (2) multi-point time history data, and (3) nonquantifiable time history data.

### PROGNOSTIC ANALYSIS OF SINGLE-POINT TIME HISTORY DATA

Single-point time history data consist of a set of data points chronologically ordered in the sequence of their occurrence. The histories of lubricant particle concentration, maximum particle size, lubricant bulk temperature, the time series of mechanical efficiency, the time recording of the amplitude of a given characteristic vibration frequency, and the shock pulse or acoustic emission intensity, etc., are of this type in a transmission. Because there is only one data point at each time instant, this type of data can be plotted conveniently and compared to determine the change of absolute levels. The rate of change also can be determined from the slope of this original plot. Greater accuracy can be obtained, however, by calculating the $\Delta x$ and $\Delta t$ at each time instant and by plotting $\Delta x/\Delta t$ versus time.

These data plots of change and rate of change, if made in appropriate time scales for sensor outputs and/or their derived parameters, can be aligned with the plots of condition specifiers (quantified inspection results) to facilitate correlative prognostic analysis. Original single-point data plots can be compared both with the condition specifiers and among themselves to determine whether or not they indicate a condition change and whether they provide an earlier or a later indication than the condition specifiers established by periodic inspection. Correlation of the percentage-wise change of a sensor output or of a derived parameter to the observed condition change provides a measure of the sensitivity of that output or parameter in identifying component mechanical condition. The rate-of-change plots of single-point data can be used in a manner similar to the original change plots. The sensitivity of the rate of change in identifying a component condition can be anticipated to be greater than that of absolute change. The variation caused by random noise, however, will be larger for rate-of-change data. Rate-of-change information can be used for short-term prognostic predictions of the time in which (1) a condition level can be expected to remain unchanged or (2) a certain change of condition will occur. For example, if a certain debris generation rate reflects a wear condition that can be measured in terms of a component dimensional change, the predictive time can be obtained by dividing the difference of two particle concentration levels corresponding to two component conditions by the respective rate-of-change value, or

$$\Delta t = \frac{\Delta x}{\left(\frac{dn}{dt}\right)} \tag{41}$$

This specific prognostic capability can be evaluated by comparing the actual time it takes the component to change from one mechanical condition level to another, based on data from the condition specifier plots or from actual observations.

Because of the single-value nature of this type of data, these relationships (based on sensor outputs and derived parameters) can be established easily by exponential smoothing if the order of the equation is small, or by use of polynomial regression methods if it is large.

In order to establish the relationship between a sensor output or derived parameter and the corresponding condition specifier, their analytical time relationships can be used. Let $x = f(t)$ represent the time relationship of the sensor output, or the derived parameter obtained either by exponential smoothing or polynomial regression. Let $c = g(t)$ represent the time relationship of the corresponding condition specifier, obtained in a similar manner. Then time "t" can be obtained by inverting the function $g(t)$, or symbolically, $t = g^{-1}(c)$. By substituting variable t into the equation representing x, the relationship between a sensor output or a derived parameter and the corresponding condition specifier becomes

$$x = f\left[g^{-1}(c)\right] \tag{42}$$

with x the sensor output or derived parameter, c the condition specifier, and t the time.

The relationship between the sensor outputs or the derived parameters themselves, or between the condition specifiers themselves, can be established in the same manner as is the relationship between a sensor output or a derived parameter and a condition specifier.

Once the time relationship of the sensor output or derived parameter itself and the cross-relationship between the sensor output or the derived parameter and the condition specifier are established, long-term prognostic prediction can be achieved by plugging into these equations the desired condition levels and solving for time t.

## ANALYSIS OF MULTIPOINT TIME HISTORY DATA

Multipoint time history data consist of a set of data curves ordered chronologically according to the time sequence of their occurrences. The debris size, the vibration frequency spectrum, and the composite exceedance distributions are of this type. Because a condition such as the dynamic condition of a mechanical component or a transmission can be represented by a curve at any time instant, the establishment of analytical relationships between sensor outputs, or derived parameters, and condition specifiers becomes less obvious and the interpretation of a mathematical analysis becomes a difficult task.

Theoretically, a time series of curves can be represented by two parametrical functions of variables, $x(t)$ and $y(t)$, instead of a single function $x(t)$ as in the previous case. The rate of change of the curve can be similarly expressed in terms of the rate of change of the two parametrical functions $x(t)$ and $y(t)$ separately, or

$$\frac{dx}{dt} \simeq \frac{\Delta x}{\Delta t} \text{ and } \frac{dy}{dt} \simeq \frac{\Delta y}{\Delta t} \tag{43}$$

The short-term prediction of time to exceedance also can be performed similarly for $x(t)$ and $y(t)$ separately, i.e.,

$$\Delta tx = \frac{\Delta x}{\left(\frac{dx}{dt}\right)} \text{ and } \Delta ty = \frac{\Delta y}{\left(\frac{dy}{dt}\right)} \tag{44}$$

Although these computations can be performed without difficulty once the mathematical functions of $x(t)$ and $y(t)$ are given, the crux of the prognostics problem clearly lies in the area of the establishment of these time relationships. Except by a cut-and-try method, there is no known way of deriving the equation systematically as is possible in the case of single-point time history data. An additional complication is the fact that in actual practice, the interpretation of a condition change is not in terms of the individual change of x and y but in terms of an overall parameter pattern change.

In the frame of a future pilot experimental program, which would be needed to establish a basis for a viable prognostic methodology, an optimal manner to analyze data of this type would be to have a semiquantitative visual observation of the three-dimensional time series performed by specialists in the relevant areas, such as particle generation, tribology, and vibration or structural dynamics. To facilitate the observation, the curves could be plotted in the same scale factor on transparent placards and stacked in order of time.

## ANALYSIS OF NONQUANTIFIABLE TIME HISTORY DATA

Nonquantifiable time history data consist of a set of observations ordered chronologically according to the time sequence of their occurrences. This type of data is nonquantifiable because the nature of the observation cannot be expressed in quantitative values. Typical examples of nonquantifiable data would be the physical shape or color of wear particles, or certain visual or NDT-type inspection findings such as crack or spall configurational characteristics. Prognostic use of that type of data would be facilitated by generation of a systematic, chronologically ordered file. Descriptive wording should be as accurate and as specific as possible; photographic records could complement this data bank. Expert personnel should subject prognostic conclusions based on this type of data to in-depth evaluation of the criteria and reasoning process employed in drawing these conclusions. Based on this evaluation, the possibility of generating a model or an algorithm of the prognostic decision processes could be investigated.

## CONCLUSIONS

The conclusions drawn from this study are:

1. The prognostic use of diagnostic data and information appears to be feasible by (1) employing recently developed or developing techniques based on component element resonance and acoustic emission phenomena, (2) interpretation of integrated system dynamic signatures in the form of composite exceedance characteristics, and (3) statistical evaluation of wear particle generation processes.

2. Component current status verification by inferential (using diagnostic sensors) and direct (using visual or nondestructive inspections) techniques represents only one tributary to a prognostic methodology. The optimum prognostic use mode of relevant data from areas such as component design theory, tribology, abnormal system physiology, and life predictive theory requires further study.

3. Prognostic interpretation of lubricant metallic particle content data should be based on a composite information index synthesizing particle size distribution, overall concentration, and material- and size-related concentrations.

4. Prognostic use of diagnostically obtained single-point, multi-point, or nonquantifiable time history data requires refinement and further development. For the purpose of prognostic prediction, the appropriate blending of existing data interpolation, extrapolation, smoothing, and feature extraction techniques requires definition.

## RECOMMENDATIONS

Listed below are programs and activities recommended to promote the art of mechanical component failure prognosis to a level allowing its use in operational form for the determination of Army aircraft current mission capability.

1. Generation of an in-depth, tribologically directed, mission load profile for major life-critical components of a representative transmission or gearbox system such as bearings and gears. This profile should include life-related factors only recently clarified in tribological research, but as yet not entering into design capability assumptions for bearings and gears.

   Recognition of the actual, tribologically significant conditions that gear components are subjected to during representative Army aircraft mission profiles has, in the past, been based only on information obtained in ad hoc programs directed at overcoming specific component design marginalities, or deficiencies, detected during the course of product-improvement programs. Obviously, the parameter selection and the special instrumentation used in such programs were tailored around the specific problems encountered with unsatisfactory components and not of a kind suitable to obtain knowledge of generally valid conditions of influence on component life, as established by recent tribological research and of major bearing on life prognosis. A program of this type should include the following activities:

   - Evaluation of propulsive gear system design practice to identify data gaps between tribological research results and life-related design-factor assumptions presently used.

   - Definition of instrumentation provisions needed to apply data acquisition methods used in tribological research to an operational gear system.

   - Definition and execution of an experimental program directed at the determination of data required to apply latest tribological concepts to component life prediction.

   - Generation of a representative mission load profile, including the data established during the experimental phase.

2. Probabilistic analysis of system life modeling based on a comprehensive number of factors contributing to life expenditure, including failure path analysis.

   To serve as reference framework for a deterministic prognostic methodology, a probabilistic system life model should be generated. This model would reveal which system components are most

critical from a system life viewpoint, either due to their inherent life potential limitations or because of their contributive importance in failure proliferation processes. Failure path analysis should be such as to clarify various current prognostic concepts as well as to establish a clear separation line between diagnosis and prognosis.

The probabilistic life model should include design life factors such as critical installation tolerances, housing distortions, expectable lubricant performance degradations, and other functional, operational and environmental conditions, such as static/dynamic overloads, limit dynamic states, lubricant temperature control irregularities, filter bypassing, and climatic extremes.

3. Definition of the prognostic potential and hardware implementation aspects of the composite exceedance technique.

The coverage capability of the composite exceedance technique in terms of the type of life-limiting phenomena recognizable by that technique should be determined. Condition-degrading phenomena not reflected in the acoustic signature of a transmission system can obviously not be covered by the technique. Technique prognostic potential is furthermore limited by the analysis time needed from the point of primary signal acquisition to end-data output. In this context the recognition of certain types of bearing failure, characterized by an extremely short time span between initial spalling and bearing element destruction, could be beyond the capability of the composite exceedance technique as implemented in an operationally suitable form.

Component isolation capability is another factor in the definition of the prognostic potential of the composite exceedance technique. The character of a change in the shape of the composite exceedance curve can give an indication of which type of system component, for instance, a bearing or gear component, is degrading. Component discrimination within a component type class may pose difficulties.

In present uses of the composite exceedance technique, deterministic vibration spectral analysis is employed upon detection of a deviation of the shape of the composite exceedance curve from empirically established limits. For operational use of the technique, this stepwise component identification procedure would be methodologically unacceptable. It is therefore necessary to explore the possibility of computationally blending the two procedures.

In order to obtain criteria for the capability, within the frame of life prognosis, of the composite exceedance technique as

compared to the more established dynamic analysis methods, recordings of transmission dynamic signatures taken during past diagnostic transmission condition determination programs should be evaluated following composite exceedance procedures.

4. Extended postmortem life expenditure analysis based on data acquired by Army maintenance and industry reliability and product quality assurance activities.

   This program would extend the insights gained for helicopter transmission and gearbox systems in the course of the Bell investigation documented in USAAVLABS Technical Report 70-66, entitled "Mode of Failure Investigations of Helicopter Transmissions," dated January 1971. To achieve this end, the proposed program should cover gear systems being part of turboprop and turboshaft engines. As the study conducted by Stewart and Hollingsworth, entitled "Development and Services Experience of the PT6A Reduction Gearbox" (Ref. No. 27), has shown, hitherto unrecognized vibrational states in a gear system can lead to cavitation damage to critical gear system components. This opens up an entirely new dimension in the recognition of life-limiting phenomena in gear systems which may play a significant role in life prognosis. Particularly, the interaction between component operational/environmental conditions and cavitation in the component life expenditure process appears to require study emphasis.

5. Experimental life test on typical components/systems using prognostic data acquisition and interpretation means.

   This program, to be conducted under closely controlled test stand conditions, would serve the purpose of determining, by simultaneous application of prognostically most promising primary data acquisition and interpretation methods as outlined in this study, the most effective method of predicting component remaining life. Sensor data acquisition should primarily be in the areas of dynamic and wear particle generation signatures. However, supporting instrumentation for parametrical, functional and environmental data should also be provided and laboratory-type lubricant physicochemical analyses periodically performed. The test specimen would be periodically subjected to detail dimensional and nondestructive material condition-verifying inspections to obtain close correlation of component condition changes toward life expenditure with signature analysis results. Computational routines needed for life prediction would be established and refined using facilities available at the test site.

6. Prognostic analysis of lubricant wear particle content.

Recent advances in the area of fluid-borne particle metrology, both instrumentationally and in the analytical systematization of wear particle generation processes (Refs. 9 through 13), summarized in this study have not been evaluated in view of their value for Army aircraft gear system life prediction. In view of these developments, concepts such as selected materials-related concentrations, near-exclusively employed in the past for life prediction (SOAP approach), can now be assigned a contributory role only in a total lubricant wear particle content monitoring approach which would also include concepts such as particle size distribution (Ref. 14), particle shape, and material-discriminating nuclear irradiation concepts. (The merit of the latter concept is particularly seen in its capability to determine the content of nonmetallic contaminations in lubricant, such as silicon, which may play a significant part in system component degradation.)

The program which would be required to establish a now possible optimum synthesis of the more recently developed wear particular content sensing and interpretative methods would consist of fitting a gear system with particle concentration, material and size-discriminating sensors (such as optical, nucleonic, conductive filtration and field-disturbance type) and, while subjecting the system to an accelerated life test, correlating the outputs of these sensors. The data obtained would then be subjected to analytical systematization similar to that established by Beerbower and Fitch for quantifying wear phenomena in components of hydraulic and fuel systems in an effort to find realistic tolerance levels for fluid-borne wear particle or contaminant content in those systems.

7. Study of hardware implementation aspects of an operational prognostic system.

This program encompasses three major areas of investigation: sensor area, data acquisition area, and data processing area. In the sensor area, sensor state-of-the-art level, performance capability in terms of phenomena coverage, environmental capability, and operational characteristics would have to be established. In the area of data acquisition, multiplexing speed and channel capacity have to be determined. Also, the ability of electronic data acquisition equipment requires definition. In the data processing area, optimum mechanizations of needed software algorithms, resulting computational complexity, computation speed, and data storage requirements have to be determined. A secondary area of this study would be directed at clarifying the partitioning of data needed for in-flight usage of prognostic procedures and usage at various levels of Army maintenance organizations.

This hardware implementation study would determine whether the finalized prognostic methods are implementable in view of technology level at the envisaged time of entering into a pre-development phase.

8. Exploratory research and development of on-line ultrasonic structural integrity system.

Evaluation of the current state of the art in the field of nondestructive test techniques has shown that, among techniques of this type, the ultrasonic pulse-echo technique could have potential as a supplementary capability of an on-line prognostic system.

Since transmission, or gearbox, structural integrity is a pre-condition for attaining design life of the dynamic components of such systems, a means of detecting loss of this integrity could be considered among prognostic desiderata. If existence of cracks, crack size, and crack growth rate could be verifiable by application of on-line ultrasonic procedures, the capability of a prognostic system would be significantly enhanced.

The effort would include design study of a snap-on type of ultrasonic sensing head permitting gearbox-external attachment as well as device-sharing among a number of installed gear systems.

Compared to the use of mechanical scanning mechanisms, an operational version of this method could employ ultrasonic beam steering by electronic methods. To provide data needed for system life prognosis, the optimum format of these data has to be established. Also, computerized digital reconstruction techniques have to be defined. A practical usage mode could consist in generating a three-dimensional shadowgraph of the structural flaw and the component, or system, in the assembled condition. This kind of data could then be reconstructed conveniently using latest, inexpensive mini-computers. Upon receiving the flaw depth-related information from the sensing head, the reconstruction could be automatically performed. Scanning would also be controlled by a digital mini-computer.

Application-wise, the system could be used as a stand-alone device furnishing indirect visual information equivalent to that obtained by actual visual inspection procedures. It would also be used, in conjunction with other prognostic processors, as the data acquisition front-end, furnishing information on structural integrity of, particularly, critical dynamic component-supporting gear system structural sections or parts.

9. Evaluation of nondestructive test methods in view of their applicability to gear system life prognosis.

The purpose of this study would be to establish areas of commonality between the objectives of nondestructive test techniques and those of gear system life prognosis. The result of the study would be the identification of NDT techniques particularly useful in view of prognostic objectives as well as the determination of type and extent of adaptations needed to employ nondestructive test techniques within the framework of a prognostic system.

In Table IV, page 40, of this report, the characteristics of NDT techniques of relevance to prognosis are summarized. Among the 18 techniques evaluated, the ultrasonic pulse-echo technique appears to have the greatest potential, and it has therefore been selected for a separate program as outlined under Item 8 of these recommendations.

The stress wave analysis technique appears to rank second in applicability potential; however, its noise contamination problem has to be solved.

Because of their simplified transmitting and receiving characteristics, the eddy current technique and the technique based on utilization of the Barkhausen effect could be considered to rank next in prognostic potential. However, the limited penetration depth obtainable by these techniques makes them not suitable for prognostic monitoring of assembled systems.

10. Study of applicability of the acoustic emission monitoring technique to gear system life prognosis.

Acoustic emission phenomena, as summarized on pages 18 through 28 of this report, are now being employed for structural integrity monitoring of static structures and also have been recognized to be applicable for dynamic machine parts, such as bearings. It appears that, in view of these recent developments, study of the applicability of the acoustic emission monitoring technique for the life prognosis of both static and dynamic gear system components would be justified. Of particular interest would be an experimental investigation directed at the observability of acoustic emission from a bearing in a quill-shaft assembly brought to failure by interruption of lubrication.

## RECOMMENDED FOLLOW-ON PRIORITIES

Because of the expected interdependency of the results, a priority assignment for the previously outlined recommended efforts appears to be in order.

The activities directed at the application of the composite exceedance technique as part of the capabilities of a prognostic system and the programs directed at exploring the prognostic potential of nondestructive test techniques, including acoustic emission, stand in a class by themselves. However, their results have a bearing on Recommendation 7, the study of hardware implementation aspects of an operational prognostic system. That study requires inputs from the preceding six programs, among which Program 5 appears to warrant first priority. The apparent high potential of the composite exceedance technique would indicate need for a priority assignment for Program 3 immediately below that of Program 5. Programs 1, 2, 4 and 6 appear to rank in a priority class between those of 3 and 5 and 8 and 9.

## LITERATURE CITED


1. Howard, P. L., SHOCK PULSE INSTRUMENTATION, paper presented at the 14th meeting of the Mechanical Failure Prevention Group (MFPG), Los Angeles, California, January 1971.

2. Minnear, J. E., et al, ADVANCED DIAGNOSTIC SENSOR APPLICATION TECHNOLOGY, The Garrett Corporation; AFAPL Technical Report 72-59, Air Force Aero Propulsion Laboratory, Wright Field, Ohio, June 1972.

3. Press Release Address by Dr. I. Fernlund of SKF on Shock Pulse Measuring System, Stockholm, April 1970.

4. Harris, D. O., A. S. Tetelman, and F. A. I. Darwish, DETECTION OF FIBER CRACKING BY ACOUSTIC EMISSION, Dunegan Research Corporation, Technical Report DRC-71-1, Livermore, California, February 1971.

5. Hutton, P. H., ACOUSTIC EMISSION ALLOWS DETECTION OF CRACK DEVELOPMENT BEFORE FAILURE OCCURS, Automotive Engineering, Volume 79, Number 8, August 1971, pp. 33-37.

6. Balderston, Harvey L., ON-BOARD INCIPIENT FAILURE DETECTION SUBSYSTEM, Boeing Report, The Boeing Company, Seattle, Washington (no date).

7. Balderston, Harvey L., and Robert E. Hine, INCIPIENT FAILURE DETECTION - THE DETECTION OF INCIPIENT FAILURE IN STRUCTURAL MATERIAL SAMPLES, Boeing Report, The Boeing Company, Seattle, Washington (no date).

8. Balderston, Harvey L., INCIPIENT FAILURE DETECTION - THE DETECTION OF INCIPIENT FAILURE IN BEARINGS, Boeing Report, The Boeing Company, Seattle, Washington (no date).

9. Beerbower, A., PARTICULATE METALS IN OILS, Proceedings of the 16th Meeting of the Mechanical Failures Prevention Group, November 2-4, 1971.

10. Cole, F. W., PARTICLE COUNT RATIONALIZATION, American Association for Contamination Control Paper, St. Louis, Missouri, 1966.

11. Fitch, E. C., Jr., REPRESENTATION OF CONTAMINATION LEVELS AND TOLERANCES, ASLE Transactions, Vol. 12, 1969, pp. 199 to 203.

12. Beerbower, A., and F. F. Tao, APPLICATIONS OF MATHEMATICAL MODELING TO CORROSIVE WEAR, Proceedings of the 15th Meeting of the Mechanical Failures Prevention Group, April 1971.

13. Beerbower, A., A CRITICAL SURVEY OF MATHEMATICAL MODELS FOR BOUNDARY LUBRICATION, ASLE Transactions, Vol. 14, 1971, pp. 90 to 104.

14. Westcott, V. C., THE NATURE OF FERROUS PARTICLES RESULTING FROM WEAR IN AIRCRAFT JET ENGINES, Proceedings of the 16th Meeting of the Mechanical Failures Prevention Group, Nov. 2 to 4, 1971.

15. Bowen, C. W., L. L. Dyson, and R. D. Walker, MODE OF FAILURE INVESTIGATIONS OF HELICOPTER TRANSMISSIONS, USAAVLABS Technical Report 70-66, U. S. Army Aviation Materiel Laboratories, Fort Eustis, Va., January 1971, AD 881 610.

16. Jantzen, E., EARLY-STAGE DETECTION OF OIL CHANGES IN AIRCRAFT ENGINES, AGARD Conference Proceedings No. 84, May 1971 (AGARD-CP-84-71).

17. Nondestructive Testing Handbook, Robert Charles McMaster, ed., New York, Ronald Press, 1959.

18. Hannary, A., NEW DETECTION DEVICE HELPS PREDICT POTENTIAL FAILURE, Product Engineering, March 1967.

19. NONDESTRUCTIVE TESTING: TRENDS AND TECHNIQUES, Proceedings of the Symposium on the Technical Status and Trends of Nondestructive Testing, Marshall Space Flight Center, 1966.

20. Whittington, K. R., and B. D. Cox, EXPERIENCE WITH ELECTRONICALLY STEERABLE TRANSDUCER ARRAY FOR ULTRASONIC TUBE TESTING, paper presented at Conference on Ultrasonics for Industry, October 20, 21, 1971.

21. Gary, D., ACOUSTIC HOLOGRAPHY, paper presented at 9th Meeting of the Mechanical Failures Prevention Group, Columbus, Ohio, November 5, 6, 1969.

22. Thompson, Brian J., and William R. Zinky, HOLOGRAPHY: A STATUS REPORT, Research/Development, July 1967.

23. Gardner, C. G., and J. R. Barton, RECENT ADVANCES IN MAGNETIC FIELD METHODS O. NONDESTRUCTIVE EVALUATION FOR AEROSPACE APPLICATIONS, paper presented at 35th Meeting, Propulsion and Energetics Panel, Advisory Group for Aerospace Research and Development, London, England, April 6-10, 1970.

24. Musil, F. J., EDDY CURRENT AND ULTRASONIC TECHNIQUES FOR INSPECTION OF LARGE PARTS, paper presented at the Symposium on Nondestructive Test Methods for the Aerospace Industry, ASTM 5th Pacific Area National Meeting, November 1965.

25. Harding, D. G., and J. C. Mack, DESIGN OF HELICOPTER TRANSMISSION FOR ON-CONDITION MAINTENANCE, Annals of Reliability and Maintainability, 1970, pp. 123 to 138.

26. Dawson, P. H., ROLLING CONTACT FATIGUE CRACK INITIATION IN A 0.3-PERCENT CARBON STEEL, Proceedings of the Institute of Mechanical Engineers, Vol. 183, Part 1, p. 75.

27. Stewart, A. W., and D. Hollingsworth, DEVELOPMENT AND SERVICE EXPERIENCE OF THE PT6A REDUCTION GEARBOX, SAE Paper No. 710433, dated May 1971.

28. Chiu, Y. P., T. E. Tallian, J. I. McCool, and J. A. Martin, A MATHEMATICAL MODEL OF SPALLING FATIGUE FAILURE IN ROLLING CONTACT, ASLE Transactions, Vol. 12, 1969.

29. Bartel, A., and E. Kuss, SOME FUNDAMENTAL EXPERIMENTS ON SURFACE DEFORMATION AND PITTING, Freiberger Forschungshefte (Research Transactions of the Freiberg Mining Academy), Vol. 35, p. 112, 1959.

30. Rothbart, H. A., MECHANICAL DESIGN AND SYSTEMS HANDBOOK, McGraw-Hill, 1967.

31. Juvinall, R. C., STRESS, STRAIN, AND STRENGTH, McGraw-Hill, 1967.

32. Lipson, C., and R. C. Juvinall, HANDBOOK OF STRESS AND STRENGTH, Macmillan, 1963.

33. Lin, Y. K., PROBABILISTIC THEORY OF STRUCTURAL DYNAMICS, McGraw-Hill, 1967.